A study on the Lévy trees and the inhomogeneous continuum random trees

Minmin Wang

▶ To cite this version:

Minmin Wang. A study on the Lévy trees and the inhomogeneous continuum random trees. General Mathematics [math.GM]. Université Pierre et Marie Curie - Paris VI, 2014. English. NNT : 2014PA066467 . tel-01087262

HAL Id: tel-01087262
https://theses.hal.science/tel-01087262

Submitted on 25 Nov 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Université Pierre et Marie Curie

Laboratoire de Probabilités et Modèles Aléatoires

École Doctorale des Sciences Mathématiques de Paris Centre

THÈSE DE DOCTORAT

Discipline : Mathématiques

présentée par

Minmin WANG

CONTRIBUTIONS À L'ÉTUDE DES ARBRES DE LÉVY ET DES ARBRES INHOMOGÈNES CONTINUS

---

co-dirigée par Nicolas BROUTIN et Thomas DUQUESNE

Soutenue le 3 décembre 2014 devant le jury composé de :

M. Romain ABRAHAM	Université d'Orléans	rapporteur
M. Nicolas CURIEN	Université Paris-Sud Orsay	examinateur
M. Thomas DUQUESNE	Université Pierre et Marie Curie	directeur
Mme Bénédicte HAAS	Université Paris-Dauphine	examinatrice
M. Zhan SHI	Université Pierre et Marie Curie	examinateur
M. Zhiying WEN	Tsinghua University	examinateur

UNIVERSITÉ P. & M. CURIE (PARIS 6),
LABORATOIRE DE PROBABILITÉS ET MODÈLES ALÉATOIRES
Case courrier 188
UMR 7599,
4 place Jussieu
75 252 Paris Cedex 05

INRIA PARIS-ROCQUENCOURT
PROJET RÉSEAUX, ALGORITHMES ET PROBABILITÉS.
Domaine de Voluceau, B. 9
Rocquencourt, BP 105
78153 Le Chesnay Cedex

UNIVERSITÉ P. & M. CURIE (PARIS 6),
ÉCOLE DOCTORALE DE SCIENCES MATHÉMATIQUES DE PARIS CENTRE,
Case courrier 290,
4 place Jussieu,
75252 Paris Cedex 05

Remerciements

Je voudrais exprimer ma plus profonde gratitude à mes directeurs Nicolas Broutin et Thomas Duquesne pour le temps qu'ils m'ont consacré, le soutien qu'ils m'ont apporté et leurs conseils avisés. Si j'ai réussi à faire des progrès constants ces dernières années, c'est grâce à leurs efforts et leurs encouragements.

Je tiens a remercier chaleureusement mes deux rapporteurs: Romain Abraham et Svante Janson. Leurs commentaires ont permis d'améliorer ce manuscrit. Ma reconnaissance va également à Nicolas Curien, Bénédicte Haas, Zhan Shi pour avoir accepté de faire partie du jury de soutenance. Mes vifs remerciements à Zhiying Wen, qui m'a initialement inspiré le désir de faire la recherche. Je suis très honorée de sa présence dans mon jury.

Au LPMA, j'ai pu bénéficier d'excellentes conditions de travail. Merci à tous les anciens et actuels membres du laboratoire pour une ambiance sympathique et agréable. Merci à Xinxin Chen, Xan Duhalde et Sarah Kaakai pour leur gentillesse et l'aide qu'ils m'ont apporté. Merci aux organisateurs et aux orateurs du GTT pour la qualité des exposés. Merci également à l'équipe administrative du laboratoire pour leur travail efficace.

Merci aux membres des équipes Algorithme et RAP à l'Inria Rocquencourt pour les excellents moments passés au cour de ces dernières années. Merci à Marie Albenque, Christine Fricker, Cécile Mailler, Philippe Robert, Henning Sulzbach et tous les participants du groupe de travail « Algorithmes et Structures Aléatoires » pour les échanges et les idées éclairantes.

Merci à l'équipe probabiliste à l'Université Paris-Sud qui m'a accueillie en tant qu'ATER pendant l'année 2013/2014, et à celle de l'Université Paris-Dauphine que je vais rejoindre cette année.

Merci à Chang yinshang, Lin shen et Wu hao pour le groupe de travail sur le processus de Lévy dont j'ai beaucoup bénéficié pendant ma thèse. Merci à tous mes amis à Paris et ailleurs. Enfin, merci à Shu, pour être toujours là, aux moments de joie ainsi qu'aux moments de frustration.

~~~~~~~~~~~~~~~~~~~~~

CONTRIBUTIONS À L'ÉTUDE DES ARBRES DE LÉVY ET DES ARBRES INHOMOGÈNES CONTINUS

~~~~~~~~~~~~~~~~~~~~~

Résumé

Nous considérons deux modèles d'arbres aléatoires continus, à savoir les arbres de Lévy et les arbres inhomogènes. Les arbres de Lévy sont limites d'échelle des arbres de Galton–Watson. Ils décrivent les structures généalogiques des processus de branchement continus en temps et en espace. La classe des arbres de Lévy est introduite par Le Gall et Le Jan (1998) comme extension de l'arbre brownien d'Aldous (1991). Nous donnons une description de la loi d'un arbre de Lévy conditionné par son diamètre, ainsi qu'une décomposition de l'arbre le long de ce diamètre, qui est décrite à l'aide d'une mesure ponctuelle de Poisson. Dans le cas particulier d'un mécanisme de branchement stable, nous caractérisons la loi jointe du diamètre et de la hauteur d'un arbre de Lévy conditionné par sa masse totale. Dans le cas brownien nous obtenons une formule explicite de cette loi jointe, ce qui permet de retrouver par un calcul direct sur l'excursion brownienne, un résultat de Szekeres (1983) et Aldous (1991) concernant la loi du diamètre. Dans les cas stables, nous obtenons également des développements asymptotiques pour les lois de la hauteur et du diamètre.

Les arbres inhomogènes sont introduits par Aldous et Pitman (2000), Camarri et Pitman (2000). Ce sont des généralisations de l'arbre brownien d'Aldous (et des arbres de Lévy). Pour un arbre inhomogène, nous étudions une fragmentation de cet arbre qui généralise celle introduite par Aldous et Pitman pour l'arbre brownien. Nous construisons un arbre généalogique de cette fragmentation. En utilisant des arguments de convergence, nous montrons qu'il y a une dualité en loi entre l'arbre initial et l'arbre généalogique de fragmentation. Pour l'arbre brownien, nous trouvons aussi une façon de reconstruire l'arbre initial à partir de l'arbre généalogique de fragmentation.

Mots-clefs : arbre brownien, excursion brownienne, continuum random tree, fonction theta de Jacobi, décomposition de Williams, arbres de Lévy, processus de Lévy, processus des hauteurs, diamètre, décomposition, développement asymptotique, lois stables, arbres inhomogènes continus, cut trees, processus de fragmentation.

~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~

CONTRIBUTION TO THE STUDY OF LÉVY TREES AND OF INHOMOGENEOUS CONTINUUM RANDOM TREES

~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~

Abstract

We consider two models of random continuous trees: Lévy trees and inhomogeneous continuum random trees. Lévy trees are scaling limits of Galton–Watson trees. They describe the genealogical structures of continuous-state branching processes. The class of Lévy trees is introduced by Le Gall and Le Jan (1998) as an extension of Aldous' notion of Brownian Continuum Random Tree (1991). For a Lévy tree, we give a description of its law conditioned to have a fixed diameter that is expressed in terms of a Poisson point measure. In the special case of a stable branching mechanism, we characterize the joint law of the diameter and the height of a Lévy tree conditioned on its total mass. From this, we deduce explicit distributions for the diameter in the Brownian case, as well as tail estimates in the general case.

Inhomogeneous continuum random trees are introduced by Aldous and Pitman (2000), Camarri and Pitman (2000). They are also generalizations of Aldous' Brownian Continuum Random Tree (and of Lévy trees). For an inhomogeneous continuum random tree, we consider a fragmentation which generalizes the one introduced by Aldous and Pitman on the Brownian tree. We construct a genealogical tree for this fragmentation. With weak limit arguments, we show that there is a duality in distribution between the initial tree and the genealogical tree. For the Brownian tree, we also present a way to reconstruct the initial tree from the genealogical tree.

Keywords: Brownian tree, Brownian excursion, continuum random tree, Jacobi theta function, Williams' decomposition, Lévy trees, Lévy process, height process, diameter, decomposition, asymptotic expansion, stable law, inhomogeneous continuum random trees, cut tree, fragmentation process.

Contents

Chapter 1

Introduction

In this PhD Thesis, we study Lévy trees introduced by Le Gall and Le Jan [83], as well as inhomogeneous continuum random trees (ICRT) defined by Aldous and Pitman [13], Camarri and Pitman [41]. This thesis contains four articles.

- [100]: HEIGHT AND DIAMETER OF BROWNIAN TREES, submitted.
- [54]: DECOMPOSITION OF LÉVY TREES ALONG THEIR DIAMETER, joint work with Thomas Duquesne (co-advisor). This paper is submitted.
- [39]: CUTTING DOWN $\boldsymbol{p}$-TREES AND INHOMOGENEOUS CONTINUUM RANDOM TREES, joint work with Nicolas Broutin (co-advisor). This paper is submitted.
- [40]: REVERSING THE CUT TREE OF THE BROWNIAN CRT, joint work with Nicolas Broutin (co-advisor). This paper is submitted.

Chapter 2 and Chapter 3 concern Lévy trees. Chapter 2 is based on the work [100], where we study the diameter of Aldous' Brownian tree. Chapter 3 is based on the joint work with Duquesne [54]: in this article, we study a decomposition of Lévy trees along their diameter and we obtain results on the total height and on the diameter of stable trees conditioned on their total mass, which generalizes several formulæ obtained in Chapter 2. Chapter 4 and Chapter 5 concern ICRTs. Chapter 4 is based on the joint work with Broutin [39], where we consider general fragmentations on ICRTs. Chapter 5 is based on the joint work with Broutin [40], where we solve a problem that arises naturally from Chapter 4.

In this chapter, we first introduce the main mathematical objects we consider; then, we present the main results whose proofs are to be found in the remaining chapters. We adopt the following notation

$$\mathbb{R}_+ := [0, \infty), \quad \mathbb{N} := \{1, 2, 3, \cdots\} \quad \text{and} \quad \mathbb{N}_0 := \{0, 1, 2, \cdots\}.$$

Unless otherwise specified, all the random variables below are defined on the same probability space denoted by

$$(\Omega, \mathcal{F}, \mathbf{P}).$$

1.1 Galton–Watson trees and Lévy trees

For details and proofs on Galton–Watson processes and Galton–Watson trees, we refer to Athreya and Ney [20], to Lyons and Peres [86] and to Neveu [92]. On the coding of trees by functions, see Le Gall and Le Jan [83] and the introduction of Duquesne and Le Gall [51].

1.1.1 Galton–Watson processes

Let $(Y_i^{(n)}, i \in \mathbb{N}, n \in \mathbb{N}_0)$ be a sequence of independent $\mathbb{N}_0$-valued variables whose common law is $\mu = (\mu(k), k \in \mathbb{N}_0)$, which is said to be the *offspring distribution*. We denote by $f_\mu(r) = \sum_{k=0}^\infty r^k \mu(k)$ the generating function of μ. A Galton–Watson process $(Z_n)_{n\in\mathbb{N}_0}$ of offspring law μ starting from $a \in \mathbb{N}_0$ can be defined in the following inductive way:

$$Z_0 = a, \qquad Z_{n+1} = \mathbf{1}_{\{Z_n \geq 1\}} \sum_{1\leq i\leq Z_n} Y_i^{(n)}, \qquad \text{for each } n \in \mathbb{N}_0. \tag{1.1}$$

The process $(Z_n)_{n\in\mathbb{N}_0}$ is a Markov chain whose transition probabilities are characterized by

$$\mathbb{E}[r^{Z_{n+1}} \mid Z_n] = f_\mu(r)^{Z_n}, \qquad \text{for all } r \in [0,1], n \in \mathbb{N}_0. \tag{1.2}$$

The Markov chain $(Z_n)_{n\in\mathbb{N}_0}$ describes a population which evolves in the following way. At generation 0, there are exactly a individuals, which are the ancestors of the population. At generation $n \in \mathbb{N}_0$, each individual independently gives birth to a random number of children according to the law μ. The generation $n+1$ consists of the children of the individuals of generation n.

Now let us consider two independent populations: one has a ancestors; the other has b ancestors. It is clear that the union of the two populations has the same distribution as a population which has $a+b$ ancestors. This is often called the *branching property* of the Galton–Watson process. More precisely, denote by $P_n(a,\cdot)$ the law of Z_n with $Z_0 = a$ as defined in (1.1). The branching property is equivalent to say that

$$\forall n, a, b \in \mathbb{N}_0 \qquad P_n(a,\cdot) * P_n(b,\cdot) = P_n(a+b,\cdot), \tag{1.3}$$

where $*$ denotes the convolution product for laws on $\mathbb{N}_0$.

Let $q = \inf\{r \in [0,1] : f_\mu(r) \leq r\}$ be the smallest fixed point of f_μ on $[0,1]$. Then it is not difficult to show that

$$\mathbf{P}\left(\exists\, n \in \mathbb{N}_0 : Z_n = 0\right) = \mathbb{E}[q^{Z_0}]. \tag{1.4}$$

On the other hand, if $\sum_{k=0}^\infty k\mu(k) \leq 1$, then $q = 1$ and it follows from (1.4) that the population becomes extinct almost surely. If $\sum_{k=0}^\infty k\mu(k) > 1$, the population has a strictly positive probability of surviving. We say that μ is *sub-critical* if $\sum_{k=0}^\infty k\mu(k) < 1$, *critical* if $\sum_{k=0}^\infty k\mu(k) = 1$ and *super-critical* if $\sum_{k=0}^\infty k\mu(k) > 1$.

In this work, we are only interested in the critical and subcritical cases.

1.1.2 Plane trees and Galton–Watson trees

In this section, we describe the genealogy of Galton–Watson processes, using Ulam's formalism as discussed in Neveu [92]. Let

$$\mathbb{U} = \bigcup_{n=0}^\infty \mathbb{N}^n \,,$$

where we make the convention that $\mathbb{N}^0 = \{\varnothing\}$. An element of $\mathbb{U}$ is then a finite sequence $u = (u_1, u_2, \cdots, u_m)$ of strictly positive integers. The length of this sequence is said to be the *generation* of u; we denote it by $|u|$. If $u = (u_1, u_2, \cdots, u_n)$ and $v = (v_1, v_2, \cdots, v_m)$ are two elements of $\mathbb{U}$, we write $uv = (u_1, u_2, \cdots, u_n, v_1, v_2, \cdots, v_m)$ for the concatenation of u and v. In particular, we have $\varnothing u = u\varnothing = u$.

A *plane tree* T is a finite subset of $\mathbb{U}$ that satisfies the following conditions:

(a) $\varnothing \in T$; $\varnothing$ is called the *root* of T;

(b) if $v \in T$ and $v = uj$ for some $u \in \mathbb{U}$ and $j \in \mathbb{N}$, then $u \in T$;

(c) for each $u \in T$ there exists a number $k_u(T) \in \mathbb{N}_0$ such that $uj \in T$ if and only if $j \leq k_u(T)$.

We denote by $\mathbb{T}_{\rm pl}$ the set of plane trees. For all $T \in \mathbb{T}_{\rm pl}$, we write $\zeta(T) := \operatorname{Card} T$, the total size of T. Let $u \in T$; the *subtree* of T stemming from u is denoted by $\operatorname{Sub}_u(T)$: namely, $\operatorname{Sub}_u(T) = \{v \in \mathbb{U} : uv \in T\}$. Note that $\operatorname{Sub}_u(T)$ is also a plane tree.

We equip $\mathbb{T}_{\rm pl}$ with the σ-algebra $\mathcal{G}$ generated by the sets $\{T \in \mathbb{T}_{\rm pl} : u \in T\}, u \in \mathbb{U}$. Let μ be a (sub)-critical probability distribution on $\mathbb{N}_0$. Neveu [92] has shown that there exists a unique probability Q_μ on $(\mathbb{T}_{\rm pl}, \mathcal{G})$ such that

(1) $Q_\mu(k_\varnothing(T) = j) = \mu(j)$, for $j \in \mathbb{N}_0$;

(2) for every $j \in \mathbb{N}$ with $\mu(j) > 0$, the subtrees $\operatorname{Sub}_1(T), \operatorname{Sub}_2(T), \cdots, \operatorname{Sub}_j(T)$ are independent under the conditional probability $Q_\mu(\cdot | k_\varnothing(T) = j)$ and their conditional distribution is Q_μ.

The distribution Q_μ is called the *law of the Galton–Watson tree with offspring law* μ. If we denote by $Z_n(T) = \operatorname{Card}\{u \in T : |u| = n\}$ the size of generation n, then $(Z_n(T), n \in \mathbb{N}_0)$ under Q_μ is a Galton–Watson process with offspring law μ starting from 1. Note that the (sub)-criticality of the offspring law μ guarantees that T is Q_μ-a.s. finite.

1.1.3 Encoding Galton–Watson trees

Let $T \in \mathbb{T}_{\rm pl}$. We associate with T two coding functions, namely the *height function* and the *contour function*.

Discrete height process. To define the height function, we observe that the lexicographic order of $\mathbb{U}$ induces a linear order on T. Let us index the vertices of T in this order: $u(0) = \varnothing, u(1), \cdots, u(\zeta(T)-1)$. Then the height function $(H_n(T), 0 \leq n \leq \zeta(T)-1)$ is defined by

$$H_n(T) := |u(n)|, \quad \forall n \in \{0, \cdots, \zeta(T)-1\}.$$

Contour process. Suppose that T is embedded in the clockwise oriented upper half plane in such a way that the root is at the origin and that each edge corresponds to a line segment of length 1. Imagine that a particle explores T from the left to the right at unit speed, starting from the root and backtracking as less as possible. Let $C_s(T)$ denote the distance between the root and the position of the particle at time s. Note that the particle returns to the root at time $2(\zeta(T)-1)$ (each edge is visited exactly twice by the particle). The function $(C_s(T), 0 \leq s \leq 2(\zeta(T)-1))$ is said to be the contour function of T (see figure 1.1). In particular, if we denote by v_k the vertex visited by the particle at time k, $k \in \{0, 1, \cdots, 2(\zeta(T)-1)\}$, then the graph distance d of T satisfies for any $0 \leq k_1 \leq k_2 \leq 2(\zeta(T)-1)$,

$$d\left(v_{k_1}, v_{k_2}\right) = C_{k_1}(T) + C_{k_2}(T) - 2 \min_{s \in [k_1, k_2]} C_s(T). \tag{1.5}$$

We shall discuss a generalization of this formula in the continuous context.

Height process and contour function of forests. We also need the notion of the height function (resp. contour function) of a forest. A *forest* is a sequence $(T_k)_{k \in \mathbb{N}}$ where $T_k \in \mathbb{T}_{\rm pl}$. The height function of the forest $(T_k)_{k \in \mathbb{N}}$ is obtained by the concatenation of the height functions of each T_k: for each $k \geq 1$, we define

$$H_n = H_{n-(\zeta(T_1)+\cdots+\zeta(T_{k-1}))}(T_k), \quad \text{if } \zeta(T_1) + \cdots + \zeta(T_{k-1}) \leq n < \zeta(T_1) + \cdots + \zeta(T_k).$$

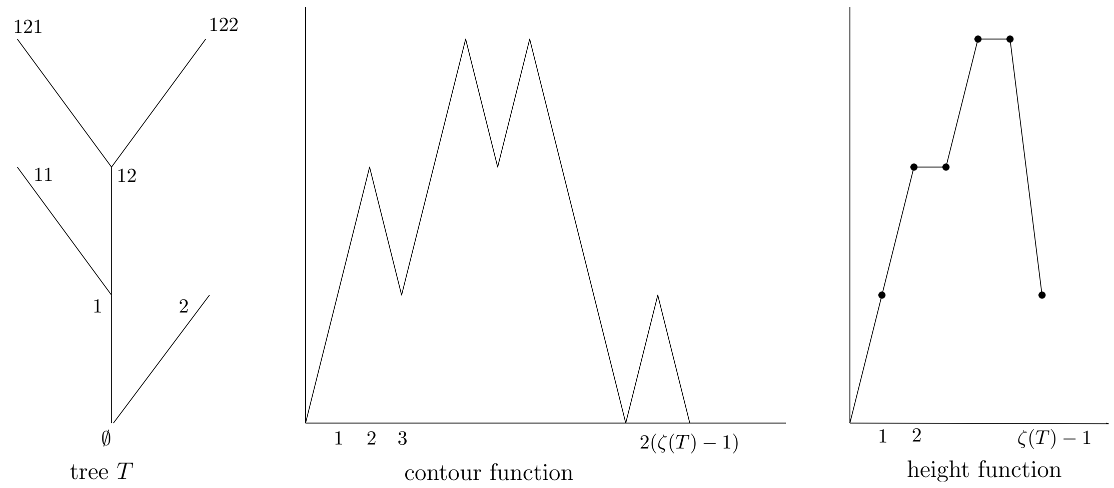

Figure 1.1

It is not difficult to see that the height function $(H_n, n \in \mathbb{N}_0)$ characterizes the forest.

To define the contour function of the forest $(T_k)_{k\in\mathbb{N}}$, one needs to extend the definitions of the contour functions $(C_s(T_k), 0 \le s \le 2\zeta(T_k) - 2)$ by setting $C_s(T_k) = 0$ for all $s \in (2\zeta(T_k) - 2, 2\zeta(T_k)]$. The contour function $(C_s, s \in \mathbb{R}_+)$ of the forest $(T_k)_{k\in\mathbb{N}}$ is then obtained as the concatenation of the extended functions $(C_s(T_k), 0 \le s \le 2\zeta(T_k))$, that is, for each $s \in \mathbb{R}_+$,

$$C_s = C_{s-(2\zeta(T_1)+\cdots+2\zeta(T_{k-1}))}(T_k), \quad \text{if } 2\zeta(T_1) + \cdots + 2\zeta(T_{k-1}) \le s < 2\zeta(T_1) + \cdots + 2\zeta(T_k).$$

This definition allows us to identify the contour function of each tree of the whole forest. In fact, the contour function can be expressed in terms of the height function. We refer to the Section 2.4 of Duquesne and Le Gall [51] for more details.

Connection with random walks. Let $(T_k)_{k\in\mathbb{N}}$ be a sequence of independently and identically distributed (i.i.d.) Galton–Watson trees. It turns out that the height function of $(T_k)_{k\in\mathbb{N}}$ is distributed as a rather simple functional of a random walk.

Lemma 1 (Le Gall and Le Jan [83])**.** *Let $(X_n, n \in \mathbb{N}_0)$ be a random walk starting from 0 whose jump distribution is given by $\nu(k) := \mu(k+1)$ for $k \in \{-1, 0, 1, \ldots\}$. For each $n \in \mathbb{N}_0$, we set*

$$H_n := \mathrm{Card}\left\{k \in \{0, 1, \cdots, n-1\} : X_k = \inf_{k \le j \le n} X_j\right\}. \tag{1.6}$$

Then the process $(H_n, n \in \mathbb{N}_0)$ has the same distribution as the height function of a forest of i.i.d. Galton–Watson trees with offspring law μ.

The relation (1.6) is the starting point of the definition of the height process in the continuous context that is recalled in Section 1.1.5.

1.1.4 Continuous-state branching processes

The continuous-state branching processes (CSBP) are the continuous analogues (in time and in state-space) of Galton–Watson processes. They are introduced by Jiřina [73] and Lamperti [78, 79] and also studied in Bingham [32]. For details and proofs, we refer to Bingham [32]; see also Kyprianou [77] and Le Gall [81].

Here we only consider the (sub-)critical CSBPs: they are $\mathbb{R}_+$-valued Feller processes whose transition kernels $(P_t(x, dy), x \in \mathbb{R}_+, t \in \mathbb{R}_+)$ satisfy the following:

$$\forall t \in \mathbb{R}_+, \forall x, x' \in \mathbb{R}_+ \quad P_t(x, \cdot) * P_t(x', \cdot) = P_t(x + x', \cdot), \tag{1.7}$$

and

$$\forall t, x \in \mathbb{R}_+ \quad \int_{\mathbb{R}_+} y P_t(x, dy) < \infty. \tag{1.8}$$

Property (1.7) is called the *continuous branching property*. It is the analogue of (1.3). The second condition (1.8) is the sub-criticality assumption.

Let us denote by $(Z_t, t \in \mathbb{R}_+)$ such a CSBP. The transition kernels of Z are characterized by their Laplace transform. More precisely, for any $s, t, \lambda \in \mathbb{R}_+$, we have

$$\mathbb{E}\left[e^{-\lambda Z_{s+t}} \middle| Z_s\right] = \exp\left(-Z_s u_t(\lambda)\right), \tag{1.9}$$

where the mapping $t \mapsto u_t(\lambda)$ is nonnegative, differentiable and where it satisfies the equation

$$u_0(\lambda) = \lambda, \quad \text{and} \quad \frac{\partial}{\partial t} u_t(\lambda) + \Psi\left(u_t(\lambda)\right) = 0, \quad t \in [0, \infty). \tag{1.10}$$

Here the function $\Psi : \mathbb{R}_+ \to \mathbb{R}_+$ is called the *branching mechanism* of the process and it has the following Lévy-Khintchine form:

$$\Psi(\lambda) = \alpha\lambda + \beta\lambda^2 + \int_{(0,\infty)} (e^{-\lambda r} - 1 + \lambda r)\pi(dr), \tag{1.11}$$

where $\alpha, \beta \geq 0$, and where the *Lévy measure* π on $(0, \infty)$ satisfies $\int_0^\infty (r \wedge r^2)\pi(dr) < \infty$. Equation (1.10) has a unique solution. This implies that the branching mechanism Ψ characterizes the law of the CSBP Z. Therefore, we can talk of CSBPs with branching mechanism Ψ.

Example. Here are three examples of branching mechanisms.

- $\Psi(\lambda) = \alpha\lambda$ with $\alpha \in \mathbb{R}_+$. The associated CSBP is the deterministic process $Z_t = Z_0 e^{-\alpha t}$.
- $\Psi(\lambda) = \lambda^2$. In this case, we have $u_t(\lambda) = \frac{\lambda}{1+t\lambda}$. The associated CSBP is the Feller diffusion process, which is the solution to the SDE: $dZ_t = \sqrt{2Z_t} dB_t$, where $(B_t, t \geq 0)$ is a standard Brownian Motion.
- (Non-Brownian stable cases) $\Psi(\lambda) = \lambda^\gamma$, with $\gamma \in (1, 2)$. Then, $\alpha = \beta = 0$ and $\pi(dr) = \frac{\gamma(\gamma-1)}{\Gamma(2-\gamma)} r^{-\gamma-1} dr$. In this case, $u_t(\lambda)$ is also explicit :

$$\forall t, \lambda \in \mathbb{R}_+, \quad u_t(\lambda) = \left((\gamma-1)t + \lambda^{-(\gamma-1)}\right)^{-\frac{1}{\gamma-1}}.$$

Let us briefly discuss the asymptotic behavior of a sub-critical CSBPs. First, note that the function $t \mapsto u(t, \lambda)$ is decreasing: an easy change of variables shows that

$$\forall t \in \mathbb{R}_+, \forall \lambda \in (0, \infty), \quad \int_{u(t,\lambda)}^{\lambda} \frac{du}{\Psi(u)} = t. \tag{1.12}$$

This entails that $\lim_{t\to\infty} u(t, \lambda) = 0$ and therefore

$$\mathbf{P}\left(\lim_{t\to 0} Z_t = 0\right) = 1. \tag{1.13}$$

In contrast to Galton–Watson processes, there are two distinct scenarios for Z to get extinct. Indeed, we can deduce from (1.12) that

$$\int^{\infty} \frac{d\lambda}{\Psi(\lambda)} < \infty \quad \Longleftrightarrow \quad \mathbf{P}(\exists\, t \in \mathbb{R}_+ : Z_t = 0) = 1. \tag{1.14}$$

This is often referred to as the *Grey condition*. When (1.14) is satisfied, we set

$$\forall t \in \mathbb{R}_+ \quad v(t) = \lim_{\lambda \to \infty} u_t(\lambda) \tag{1.15}$$

and (1.12) implies that $v : (0, \infty) \to (0, \infty)$ is a continuous bijection such that

$$\int_{v(t)}^{\infty} \frac{d\lambda}{\Psi(\lambda)} = t \quad \text{and} \quad \mathbf{P}(Z_t = 0 | Z_0) = e^{-Z_0 v(t)}, \quad \forall t \in (0, \infty)\ . \tag{1.16}$$

If the Grey condition (1.14) is not satisfied, then $\mathbf{P}$-a.s. for any $t \in \mathbb{R}_+$, $Z_t > 0$.

1.1.5 Genealogy of continuous branching processes: the height process

In this part, we recall the construction of the continuous height process due to Le Gall and Le Jan [83]. See also Chapter 1 of Duquesne and Le Gall [51].

The height process. Let Ψ be the branching mechanism as defined in (1.11). Let $X = (X_t, t \in \mathbb{R}_+)$ be a spectrally positive Lévy process starting from 0 and whose Laplace exponent is given by Ψ : namely,

$$\forall t, \lambda \in \mathbb{R}_+, \qquad \mathbb{E}[e^{-\lambda X_t}] = \exp(t\Psi(\lambda))\ . \tag{1.17}$$

The process X plays a similar role as the random walk $(X_n, n \in \mathbb{N}_0)$ in Lemma 1. If Ψ satisfies the Grey condition (1.14), then Le Gall and Le Jan [83] have proved that there exists a continuous process $H = (H_t, t \in \mathbb{R}_+)$ such that for all $t \in \mathbb{R}_+$, the following limit holds true in $\mathbf{P}$-probability:

$$H_t = \lim_{\epsilon \to 0} \frac{1}{\epsilon} \int_0^t ds \mathbf{1}_{\{I_t^s < X_s < I_t^s + \epsilon\}}, \tag{1.18}$$

where we have set $I_t^s = \inf_{s \le r \le t} X_r$ for all $0 \le s \le t$. This definition is a continuous analogue of (1.6) and the process H is called the Ψ*-height process*.

In the case where $\Psi(\lambda) = \lambda^2$, the Lévy process X is a multiple of the Brownian motion and H has the same distribution as a reflected (non-standard) Brownian motion.

Ray-Knight theorem for the height process. Let us point out that, in general H, is not a Markov process. It is however possible to define the local times $(L_t^a)_{a,t \in \mathbb{R}_+}$ of H. Namely, $(a, t) \mapsto L_t^a$ is measurable, for all $a \in \mathbb{R}_+$, $\mathbf{P}$-a.s. $t \mapsto L_t^a$ is non-decreasing and

$$\mathbf{P}\text{-a.s.} \quad L_t^a = \lim_{\epsilon \to 0} \frac{1}{\epsilon} \int_0^t \mathbf{1}_{\{a < H_s < a + \epsilon\}} ds. \tag{1.19}$$

We refer to Duquesne and Le Gall [51], Proposition 1.3.3. for more details.

For all $x \in \mathbb{R}_+$, we denote $\tau_x = \inf\{t \ge 0 : X_t = -x\}$ the hitting time of X at level $-x$. Theorem 1.4.1 of [51] generalizes the Ray–Knight Theorem to a general branching mechanism Ψ:

$$\left(L_{\tau_x}^a, a \ge 0\right) \text{ is a CSBP of branching mechanism } \Psi \text{ starting from } x. \tag{1.20}$$

Informally, (1.20) says that the population at level a forms a continuous branching process. This property is expected from the continuous height process as its discrete counterpart enjoys a similar one.

Limit theorems for the height process. The following convergence results also show that H represents the genealogy of CSBPs. For each $p \in \mathbb{N}$, let $\mu_p = \{\mu_p(k), k \in \mathbb{N}_0\}$ be a (sub)-critical probability distribution on $\mathbb{N}_0$. Let $Z^p = (Z^p_k, k \in \mathbb{N}_0)$ be a Galton–Watson process with offspring law μ_p starting from p and let $X^p = (X^p_k, k \in \mathbb{N}_0)$ be a random walk with jump distribution $\nu_p = (\nu_p(k), k \in \{-1, 0, \cdots\})$ where $\nu_p(k) = \mu_p(k+1)$. Recall that X is a Lévy process of Laplace exponent Ψ and let Z be a CSBP of branching mechanism Ψ starting from 1. Let $(u_p)_{p\in\mathbb{N}}$ be a nondecreasing sequence of positive integers converging to ∞. Then a version of Grimvall's Theorem [63] (see also Theorem 2.1.1 in [51]) says that

$$\left(p^{-1} Z^p_{\lfloor u_p t\rfloor}, t \geq 0\right) \xrightarrow[p\to\infty]{d} (Z_t, t \geq 0) \quad \text{iff} \quad \left(p^{-1} X^p_{\lfloor p u_p t\rfloor}, t \geq 0\right) \xrightarrow[p\to\infty]{d} (X_t, t \geq 0), \tag{1.21}$$

where $\lfloor\cdot\rfloor$ stands for the the integer part function and where $\xrightarrow{d}$ means convergence in distribution on the space $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$ of càdlàg functions equipped with Skorokhod topology.

For each $p \in \mathbb{N}$, let $H^p = (H^p_n, n \in \mathbb{N}_0)$ and $C^p = (C^p_s, s \in \mathbb{R}_+)$ be the respective height function and contour function of a forest of i.i.d. Galton–Watson trees with offspring law μ_p. We assume that H^p is related to the random walk X^p by (1.6). Moreover, we assume that the number of individuals at generation n of the first p trees in this forest has the same distribution as Z^p_n. We then also assume that

$$\forall \delta > 0 \qquad \liminf_{p\to\infty} \mathbf{P}(Z^p_{\lfloor \delta u_p\rfloor} = 0) > 0 \,.$$

Then Corollary 2.5.1 of [51] proves the following convergence in distribution in $\mathbb{D}(\mathbb{R}_+, \mathbb{R}^3)$:

$$\left(p^{-1} X^p_{\lfloor p u_p t\rfloor}, u_p^{-1} H^p_{\lfloor p u_p t\rfloor}, u_p^{-1} C^p_{2 p u_p t}; t \geq 0\right) \xrightarrow[p\to\infty]{d} (X_t, H_t, H_t; t \geq 0), \tag{1.22}$$

where H is the height process associated with X (see (1.18)). Furthermore, jointly with (1.22), the following convergence in distribution holds in $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$:

$$\left(p^{-1} Z^p_{\lfloor u_p a\rfloor}, a \geq 0\right) \xrightarrow[p\to\infty]{d} \left(L^a_{\tau_1}, a \geq 0\right), \tag{1.23}$$

The excursion measure of the height process. The convergence in (1.22) shows that the process $(H_t, t \geq 0)$ is the scaling limit of the height functions (resp. contour functions) of a forest of i.i.d. Galton–Watson trees. One natural question is whether we can interpret this convergence in terms of trees: this will be the subject of the next section. A related question is how to "extract" a single tree from the whole forest encoded by H. For this, we need the excursion measure of H that is introduced here.

Observe that (1.14) entails that

$$\text{either} \quad \beta > 0 \quad \text{or} \quad \int_{(0,1)} r\, \pi(dr) = \infty \,, \tag{1.24}$$

which is equivalent for the Lévy process X to have unbounded variation sample paths. Let $I_t = \inf_{s\leq t} X_s$ be the infimum process of X. It is well-known that $X - I$ is a Markov process and that 0 is regular for itself with respect to this Markov process (see Bertoin [22] Chapter VII). Moreover, $-I$ is a local time of $X - I$ at level 0. We write N for the associated excursion measure. Denote by $(g_i, d_i), i \in \mathcal{I}$, the excursion intervals of $X - I$ above 0, and by $X^i_\cdot = X_{(g_i+\cdot)\wedge d_i} - I_{g_i}$, $i \in \mathcal{I}$, the corresponding excursions. We remark that according to (1.18), if $t \in (g_i, d_i)$, then the value of H_t depends only on X^i. For each $i \in \mathcal{I}$, we set $H^i_\cdot = H_{(g_i+\cdot)\wedge d_i}$. It follows from the above remark that H^i is a functional of the excursion of $X - I$ over (g_i, d_i). Let us denote by $\mathbf{N}$ the image of N by this functional. Then the point measure

$$\sum_{i\in\mathcal{I}} \delta_{(-I_{g_i}, H^i)} \tag{1.25}$$

is distributed as a Poisson point measure on $\mathbb{R}_+ \times \mathbb{C}(\mathbb{R}_+, \mathbb{R})$ with intensity measure $dt\,\mathbf{N}$.

Note that X and H share the same lifetime under $\mathbf{N}$, which we denote by ζ. Then it is classical from the fluctuation theory that $\zeta < \infty$, $\mathbf{N}$-a.e. and that for $\lambda > 0$,

$$\mathbf{N}\left(1 - e^{-\lambda\zeta}\right) = \Psi^{-1}(\lambda), \tag{1.26}$$

where Ψ^{-1} stands for the inverse function of Ψ.

Notation. We shall denote by H the canonical process on $\mathbb{C}(\mathbb{R}_+, \mathbb{R})$.

1.1.6 Lévy trees

Real trees. Real trees have been studied for a long time for algebraic and geometric purposes (see for example Dress, Moulton, and Terhalle [47]). Since the work of Evans, Pitman, and Winter [59], real trees are widely adopted for the study of random branching structures. More details and proofs on real trees and Gromov–Hausdorff distance can be found in the book of Evans [57]. The Gromov–Prokhorov distance is introduced in Greven, Pfaffelhuber, and Winter [62], see also the Chapter $3\frac{1}{2}$ of Gromov [64].

A metric space (T, d) is called a *real tree* if the following two properties hold for every $u, v \in T$.

(i) There is a unique isometric map $q_{u,v}$ from $[0, d(u, v)]$ into T such that $q_{u,v}(0) = u$ and $q_{u,v}(d(u, v)) = v$. In this case, we denote by $[\![u, v]\!]$ the image of $[0, d(u, v)]$ by $q_{u,v}$.

(ii) If q is a continuous injective map from $[0, 1]$ into T, then we have

$$q([0, 1]) = [\![q(0), q(1)]\!].$$

Among connected metric spaces, real trees are characterized by the so-called *four points inequality*: let (T, d) be a connected metric space; then (T, d) is a real tree iff for any $\sigma_1, \sigma_2, \sigma_3, \sigma_4 \in T$, we have

$$d(\sigma_1, \sigma_2) + d(\sigma_3, \sigma_4) \leq \big(d(\sigma_1, \sigma_3) + d(\sigma_2, \sigma_4)\big) \vee \big(d(\sigma_1, \sigma_4) + d(\sigma_2, \sigma_3)\big). \tag{1.27}$$

See Evans [57] for more details.

A *rooted* real tree is a real tree (T, d) with a distinguished point r called the *root*. Let (T, d, r) be a rooted real tree. For $u \in T$, the *degree* of u in T, denoted by $\deg(u, T)$, is the number of connected components of $T \setminus \{u\}$. It is possible that $\deg(u) = \infty$. We also denote by

$$\mathrm{Lf}(T) = \{u \in T : \deg(u, T) = 1\} \quad \text{and} \quad \mathrm{Br}(T) = \{u \in T : \deg(u, T) \geq 3\} \tag{1.28}$$

the set of the *leaves* and the set of *branch points* of T. The skeleton of T is the complementary set of $\mathrm{Lf}(T)$ in T, that is denoted by $\mathrm{Sk}(T)$:

$$\mathrm{Sk}(T) := T \setminus \mathrm{Lf}(T)\ .$$

Gromov–Hausdorff distance. Two rooted real trees (T, d, r) and (T', d', r') are said to be isometric if there exists an isometry $f : T \to T'$ satisfying $f(r) = r'$. We denote by $\mathbb{T}_c$ the set of pointed isometry classes of rooted compact real trees. We equip $\mathbb{T}_c$ with the pointed Gromov–Hausdorff metric, which is defined as follows: If (E, δ) is a metric space, we write δ_{H} for the usual Hausdorff metric between the compact subsets of E. If $(T, d, r), (T', d', r')$ are two rooted compact real trees, the distance between them is given by

$$\delta_{\mathrm{GH}}(T, T') = \inf\Big(\delta_{\mathrm{H}}\left(\phi(T), \varphi(T')\right) \vee \delta\big(\phi(r), \varphi(r')\big)\Big)$$

where the infimum is over all the isometric embeddings $\phi : T \to E$ and $\varphi : T' \to E$ of T and T' into some common metric space (E, δ). One readily checks that $\delta_{\mathrm{GH}}(T, T')$ only depends on the equivalence classes of T and T'. Indeed, δ_{GH} defines a metric on $\mathbb{T}_c$. Moreover, the metric space $(\mathbb{T}_c, \delta_{\mathrm{GH}})$ is complete and separable. See Evans, Pitman, and Winter [59] and Gromov [64]; see also Evans [57].

Gromov–Prokhorov distance. We say a triple (T, d, μ) is a *measured real tree* if (T, d) is a separable and complete real tree and if μ is a Borel probability measure of T. Two measured real trees (T, d, μ) and (T', d', μ') are said to be *weakly isometric* if there exists an isometry ϕ between the supports of μ on T and of μ' on T' such that μ' is the push-forward measure of μ by ϕ, which is denoted by $(\phi)_*\mu$. We denote by $\mathbb{T}_w$ the set of weak isometry classes of measures real trees.

If (E, δ) is a metric space, we denote by δ_{P} the Prokhorov distance on the set of Borel probability measures on E. Let (T, d, μ), (T', d', μ') be two measured real trees; then the Gromov–Prokhorov distance between them is defined by

$$\delta_{\mathrm{GP}}(T, T') = \inf \delta_{\mathrm{P}}(\phi_*\mu, \psi_*\mu'),$$

where the infimum is taken over all spaces E and all isometries $\phi : \mathrm{supp}(\mu) \to E$ and $\psi : \mathrm{supp}(\mu') \to E$. Note that the definition of δ_{GP} depends only on the weak isometry classes of T and T'. Moreover, δ_{GP} induces a metric on $\mathbb{T}_w$. It has been shown by Greven, Pfaffelhuber, and Winter [62] that $(\mathbb{T}_w, \delta_{\mathrm{GP}})$ is separable and complete.

The coding of real trees by excursions. Recall that $H = (H_t)_{t\geq 0}$ stands for the canonical process on $\mathbb{C}(\mathbb{R}_+, \mathbb{R}_+)$. First assume that H has a compact support, that $H_0 = 0$ and that H is distinct from the null function: we call such a function a *coding function* and we then set $\zeta_H = \sup\{t > 0 : H_t > 0\}$ that is called the *lifetime* of the coding function H. Note that $\zeta_H \in (0, \infty)$. Analogously to (1.5), we set for any $s, t \in [0, \zeta]$,

$$d_H(s, t) := H_s + H_t - 2 \inf_{s\wedge t \leq u \leq s\vee t} H_u. \tag{1.29}$$

It is not difficult to see that d_H is a pseudo-metric on $[0, \zeta_H]$. We associate with it an equivalence relation $\sim_H$, which is defined by: $s \sim_H t$ whenever $d_H(s, t) = 0$. We define

$$(\mathcal{T}_H, d_H) = ([0, \zeta_H]/{\sim_H}, d_H).$$

Let $p_H : [0, \zeta_H] \to \mathcal{T}_H$ be the canonical projection. It is clearly a continuous mapping. Thus $(\mathcal{T}_H, d_H)$ is a connected compact metric space. Moreover, it is easy to check that d_H satisfies the four points inequality. Therefore, $(\mathcal{T}_H, d_H)$ is a compact real tree. We define $\rho_H := p_H(0)$ as the *root* of $\mathcal{T}_H$.

There are two additional structures on $\mathcal{T}_H$ that are useful to us. First, the *mass measure* $\mathbf{m}_H$ of $\mathcal{T}_H$ is defined to be the pushforward measure of the Lebesgue measure on $[0, \zeta_H]$ induced by p_H; namely, for any Borel measurable function $f : \mathcal{T}_H \to \mathbb{R}_+$,

$$\int_{\mathcal{T}_H} f(\sigma)\, \mathbf{m}_H(d\sigma) = \int_0^{\zeta_H} f(p_H(t))\, dt\ . \tag{1.30}$$

Note that

$$\mathbf{m}_H(\mathcal{T}_H) = \zeta_H\ .$$

The coding function H also induces a *linear order* $\leq_H$ on $\mathcal{T}_H$ that is inherited from that of $[0, \zeta_H]$: namely for any $\sigma_1, \sigma_2 \in \mathcal{T}_H$,

$$\sigma_1 \leq_H \sigma_2 \quad \Longleftrightarrow \quad \inf\{t \in [0, \zeta_H] : p_H(t) = \sigma_1\} \leq \inf\{t \in [0, \zeta_H] : p_H(t) = \sigma_2\}\ . \tag{1.31}$$

Roughly speaking, the function H is completely characterized by $(\mathcal{T}_H, d_H, \rho_H, \mathbf{m}_H, \leq_H)$: see Duquesne [50] for more detail about the coding of real trees by functions.

Lévy trees. Observe that H is $\mathbf{N}$-a.e. a coding function as defined above. Duquesne and Le Gall [52] then define the Ψ-*Lévy tree* as the real tree coded by H under $\mathbf{N}$.

Convention. When there there is no risk of confusion, we simply write

$$(\mathcal{T}, d, \rho, \mathbf{m}, \leq, p) := (\mathcal{T}_H, d_H, \rho_H, \mathbf{m}_H, \leq_H, p_H)$$

when H is considered under $\mathbf{N}$. □

Recall that $\mathrm{Lf}(\mathcal{T})$ stands for the set of leaves of $\mathcal{T}$. Then the mass measure has the following properties:

$$\mathbf{N}\text{-a.e. } \mathbf{m} \text{ is diffuse and } \mathbf{m}(\mathcal{T} \setminus \mathrm{Lf}(\mathcal{T})) = 0. \tag{1.32}$$

The Ψ-Lévy tree $(\mathcal{T}, d, \rho, \mathbf{m})$ is therefore a continuum tree according to the definition of Aldous [8]. Moreover, it is proved in Duquesne and Le Gall [52] that

$$\mathbf{N}\text{-a.e.} \qquad \forall \sigma \in \mathcal{T}, \quad \deg(\sigma, \mathcal{T}) \in \{1, 2, 3, \infty\} \,, \tag{1.33}$$

and there exist branch points of degree 3 if and only if $\beta > 0$ in (1.11). Roughly speaking, infinite branch points are due to the jumps of the underlying Lévy process. See Duquesne and Le Gall [52] for more details.

1.1.7 Stable trees

Here we consider a special class of Lévy trees, namely the class of stable trees. As we will recall below, the scaling property of underlined Lévy process enables us to define a Lévy tree conditioned on their total mass. It turns out that the Brownian Continuum Random Tree (Brownian CRT) is a special case of these Lévy trees conditioned on the total mass. The Brownian CRT was first introduced in Aldous [8], as scaling limits of discrete trees. The coding of Brownian CRT by normalized Brownian excursion is discussed by Le Gall [80] and by Aldous [10].

In this part, we fix $\gamma \in (1, 2]$ and

$$\forall \lambda \in \mathbb{R}_+, \qquad \Psi(\lambda) = \lambda^\gamma \,.$$

Note that in this case the condition (1.14) is always satisfied. The Lévy process X under $\mathbf{P}$ enjoys the following scaling property: for all $r \in (0, \infty)$, $(r^{-\frac{1}{\gamma}} X_{rt})_{t \geq 0}$ has the same law as X. This entails by (1.18) that under $\mathbf{P}$, $(r^{-\frac{\gamma-1}{\gamma}} H_{rt})_{t \geq 0}$ has the same law as H and the Poisson decomposition (1.25) implies the following:

$$\left(r^{-\frac{\gamma-1}{\gamma}} H_{rt}\right)_{t \geq 0} \quad \text{under} \quad r^{\frac{1}{\gamma}} \mathbf{N} \quad \overset{d}{=} \quad H \quad \text{under} \quad \mathbf{N} \,. \tag{1.34}$$

On the other hand, we derive from (1.26) that

$$\mathbf{N}(\zeta \in dr) = p_\gamma(r)\, dr \,, \quad \text{where} \quad p_\gamma(r) = c_\gamma r^{-1-\frac{1}{\gamma}} \quad \text{with} \quad 1/c_\gamma = \gamma \Gamma_e\left(\tfrac{\gamma-1}{\gamma}\right) \,. \tag{1.35}$$

Here Γ_e stands for Euler's Gamma function. By (1.34), there exists a family of laws on $\mathbb{C}(\mathbb{R}_+, \mathbb{R}_+)$ denoted by $\mathbf{N}(\,\cdot\,|\,\zeta = r)$, $r \in (0, \infty)$, such that

- the mapping $r \mapsto \mathbf{N}(\,\cdot\,|\,\zeta = r)$ is weakly continous on $\mathbb{C}(\mathbb{R}_+, \mathbb{R}_+)$,
- $\mathbf{N}(\,\cdot\,|\,\zeta = r)$-a.s. $\zeta = r$,
- we have
$$\mathbf{N} = \int_0^\infty \mathbf{N}(\,\cdot\,|\,\zeta = r)\, \mathbf{N}(\zeta \in dr) \,. \tag{1.36}$$

Moreover, by (1.34), $\big(r^{-\frac{\gamma-1}{\gamma}} H_{rt}\big)_{t\geq 0}$ under $\mathbf{N}(\,\cdot\,|\,\zeta = r)$ has the same law as H under $\mathbf{N}(\,\cdot\,|\,\zeta = 1)$. We call $\mathbf{N}(\,\cdot\,|\,\zeta = 1)$ the *normalized law of the γ-stable height process* and to simplify notation we set

$$\mathbf{N}_{\mathrm{nr}} := \mathbf{N}(\,\cdot\,|\,\zeta = 1)\,. \tag{1.37}$$

Thus, for all measurable functions $F : \mathbb{C}(\mathbb{R}_+, \mathbb{R}_+) \to \mathbb{R}_+$,

$$\mathbf{N}\big[F(H)\big] = c_\gamma \int_0^\infty dr\, r^{-1-\frac{1}{\gamma}}\, \mathbf{N}_{\mathrm{nr}}\Big[F\Big(\big(r^{\frac{\gamma-1}{\gamma}} H_{t/r}\big)_{t\geq 0}\Big)\Big]\,. \tag{1.38}$$

The Brownian case. When $\gamma = 2$, the height process H under $\mathbf{N}_{\mathrm{nr}}$ is distributed as $\sqrt{2}\mathcal{E}$, where $\mathcal{E} = (\mathcal{E}_s, 0 \leq s \leq 1)$ is the normalized positive Ito excursion of standard Brownian Motion (see for instance Revuz and Yor [97], Chapter XII for a definition).

Let μ be an offspring distribution on $\mathbb{N}_0$ that is assumed to be critical and to have finite variance σ^2. We also assume that μ is aperiodic. For all sufficiently large $n \in \mathbb{N}$, let T_n be a random plane tree whose law is $Q_\mu(dT\,|\,\zeta(T) = n)$: namely T_n is a μ-Galton-Watson plane tree conditioned to have n vertices. Recall that $(C_s(T_n), 0 \leq s \leq 2n)$ stands for the contour function of T_n as defined in Section 1.1.3. Then, Aldous [10] shows that as $n \to \infty$, the following limit holds true in distribution on $\mathbb{C}(\mathbb{R}_+, \mathbb{R}_+)$:

$$\left(\frac{\sigma}{\sqrt{n}} C_{2ns}(T_n); 0 \leq s \leq 1\right) \xrightarrow{d} (2\mathcal{E}_s\,; 0 \leq s \leq 1)\,. \tag{1.39}$$

Because of this, Aldous has defined the Brownian CRT $\mathcal{T}^{br}$ to be the real tree encoded by $2\mathcal{E}$ in the sense of (1.29) (equivalently, we can take $\Psi(\lambda) = 2\lambda^2$). Then (1.39) says that $\mathcal{T}^{br}$ is the scaling limit of conditioned Galton–Watson trees. The convergence in (1.39) has been extended by Duquesne [49] to the case where the offspring law μ is critical and is in the domain of attraction of the γ-stable law. See also Kortchemski [76] for scaling limits of Galton-Watson tree conditioned to have n leaves.

Stable trees often appear in the study of self-similar fragmentations, see for example Bertoin [25], Goldschmidt and Haas [61], Haas and Miermont [65], Miermont [89, 90]. An important example of these self-similar fragmentations is the one studied in Aldous and Pitman [11], which describes the evolution of the mass partitions of the Brownian CRT where partitions are induced by a Poisson point process. The proper definition is left to the next section, where we introduce a more general fragmentation. A recent work of Bertoin and Miermont [30] shows that the genealogical tree of Aldous–Pitman's fragmentation, equipped with a suitable metric, is also distributed as the Brownian CRT. This identity in distribution is the starting point of the problem considered in Chapter 5.

1.2 Birthday trees and inhomogeneous continuum random trees

1.2.1 Birthday trees and the Aldous–Broder Algorithm

Trees are also important objects in the graph theory, where they are usually unordered but labelled. A classical model of random (graph) trees is the Cayley tree, which is a uniformly random labelled tree of given "size". The model of birthday trees that we are about to introduce is a generalization of the Cayley tree.

Let us first recall the notion of (graph) tree: a tree t is a connected graph without cycles. A *rooted tree* is a tree with a distinguished vertex, called the *root*. Recall that an edge of t is a set of the form $\{u, v\}$ where u, v are vertices of t. When the tree is rooted, we think of the edges as pointing to the root. Then the *in-degree* of a vertex v, denoted by $k_v(t)$, is the number of the edges that can be written as $\{u, v\}$ with some vertex u such that the direction of the edge is from u to v.

Let $n \geq 1$ be some natural number. We denote by $\mathbf{T}_n$ the set of rooted trees with vertex set $[n] := \{1, 2, \cdots, n\}$. A well-known fact is that $\mathbf{T}_n$ has exactly n^{n-1} elements. We equip $\mathbf{T}_n$ with the discrete σ-algebra (the power set). Then a *Cayley tree* of size n is a random tree distributed according to the uniform distribution on $\mathbf{T}_n$. Let us note that there is a correspondence between a Cayley tree and a conditioned Galton–Watson tree with Poisson offspring distribution: if we assign labels to the nodes of the Galton–Watson tree using a uniform permutation and forget the order, we obtain a Cayley tree.

The class of birthday trees has been introduced in Camarri and Pitman [41] for the study of the general birthday problems. In a Cayley tree, all nodes behave in the same way, and the birthday trees generalize this by introducing some inhomogeneity which, as we will see later, entails interesting asymptotic behaviors when the sizes of the trees go to infinity.

Let us now be more specific. Let $\boldsymbol{p}$ be a probability measure on $[n]$ and let us write $p_i = \boldsymbol{p}(i)$ for $i \in [n]$. To exclude degenerate cases, we assume that $p_i > 0$ for each $i \in [n]$. The following so-called *Aldous–Broder Algorithm* extracts a tree from the trace of a random walk on the complete graph.

Algorithm 2 (Aldous [18], Broder [37])**.** *Let $(Y_k)_{k\geq 0}$ be a sequence of independent random variables whose common law is $\boldsymbol{p}$. Let T be the (random) graph on $[n]$ which is rooted at Y_0 and has the following edge set*

$$\{\{Y_k, Y_{k+1}\} : Y_{k+1} \notin \{Y_0, Y_1, \cdots, Y_k\}, k \geq 1\}.$$

Here is a mental picture. Take a pencil and a piece of paper, and draw T as follows.

> Start at vertex Y_0. At step $k-1$, the pencil is at vertex Y_{k-1}. If Y_k has not appeared previously, we add the new vertex Y_k and draw an edge between Y_{k-1} and Y_k; otherwise move the pencil to Y_k without drawing an edge.

Note that the random walk $(Y_k)_{k\geq 0}$ eventually visits each vertex with probability one because $p_i > 0$ for any $i \in [n]$. We observe that the edge $\{Y_k, Y_{k+1}\}$ is added only if Y_{k+1} has not been seen at time k. In consequence, this forbids the existence of a cycle in T. It follows easily that T is a connected graph without cycles, hence a tree on $[n]$. Furthermore, T has the following distribution $\pi^{(\boldsymbol{p})}$:

$$\pi^{(\boldsymbol{p})}(t) := \prod_{i\in[n]} p_i^{k_i(t)}, \qquad \text{for each } t \in \mathbf{T}_n. \tag{1.40}$$

To see why this holds, one can follow the argument of the Markov chain tree theorem (see for example Anantharam and Tsoucas [19]) and deduce that for $t \in \mathbf{T}_n$, the probability $\mathbf{P}(T = t)$ is proportional to $\pi^{(\boldsymbol{p})}(t)$. Thanks to Cayley's multinomial formula (Cayley [42], see also Rényi [96]), we have

$$\sum_{t\in\mathbf{T}_n} \pi^{(\boldsymbol{p})}(t) = \sum_{t\in\mathbf{T}_n} \prod_{i\in[n]} p_i^{k_i(t)} = \left(\sum_{i\in[n]} p_i\right)^{\sum_{i\in[n]} k_i(t)} = 1.$$

Thus, $\pi^{(\boldsymbol{p})}$ is indeed a probability measure on $\mathbf{T}_n$. In the case where $p_i = 1/n$ for each i, we have $\pi^{(\boldsymbol{p})}(t) = n^{1-n}$ for any $t \in \mathbf{T}_n$. Then the above identity echoes our previous statement that $\mathrm{Card}\, \mathbf{T}_n = n^{n-1}$. In this case T is clearly the Cayley tree.

In what follows, we often refer to a random tree with distribution given in (1.40) as a $\boldsymbol{p}$*-tree*. However, when the probability measure $\boldsymbol{p}$ is not specified, we prefer to use the alternative terminology *birthday tree*.

Let T be a $\boldsymbol{p}$-tree. Its root has distribution $\boldsymbol{p}$. This fact is clear from the construction of T by Algorithm 2, since T is rooted at Y_0. Furthermore, the $\boldsymbol{p}$-tree T enjoys an *invariance by re-rooting* property, which plays an important role in the cutting problem considered in Chapter 4. More precisely, if V is an independent vertex with distribution $\boldsymbol{p}$, let T^V denote the tree obtained by re-rooting T at V. Then it can be directly verified from (1.40) that T^V is still a $\boldsymbol{p}$-tree.

1.2.2 Inhomogeneous continuum random trees and line-breaking construction

The inhomogeneous continuum random tree (ICRT) arose as weak limits of birthday trees in Camarri and Pitman [41] and Aldous and Pitman [13]. Later, a profound study of its height process has been carried out by Aldous, Miermont, and Pitman [15]. The ICRTs are closely related to other mathematical objects. For example, Aldous and Pitman [13] use ICRTs to construct the general additive coalescent. ICRTs also appear in the weak limits of random $\boldsymbol{p}$-mappings. See Aldous, Miermont, and Pitman [16].

The *parameter space* $\boldsymbol{\Theta}$ of ICRT is the set of real-valued sequences $\boldsymbol{\theta} = (\theta_0, \theta_1, \theta_2, \cdots)$ satisfying the following conditions: $\theta_1 \geq \theta_2 \geq \theta_3 \cdots \geq 0$, $\theta_0 \geq 0$, $\sum_{i\geq 0} \theta_i^2 = 1$, and either $\theta_0 > 0$ or $\sum_{i\geq 1} \theta_i = \infty$. For each $\boldsymbol{\theta} \in \boldsymbol{\Theta}$, we can define a real tree using the following *line-breaking construction* [13, 41]. This construction can be seen as the scaling limit of Algorithm 2. In outline, it consists of cutting the half-line $[0, \infty)$ into finite-length segments, reassembling the segments as "branches" of a tree, and then completing the metric space thereby obtained. The details are as follows.

- If $\theta_0 > 0$, let $\mathrm{P}_0 = \{(u_j, v_j), j \in \mathbb{N}\}$ be a Poisson point process on the first octant $\{(x, y) : 0 \leq y \leq x\}$ of intensity measure $\theta_0^2 dxdy$, enumerated in such a way that $u_1 < u_2 < u_3 < \cdots$.

- For every $i \geq 1$ such that $\theta_i > 0$, let $\mathrm{P}_i = \{\xi_{i,j}, j \in \mathbb{N}\}$ be a homogeneous Poisson process on $[0, \infty)$ of intensity θ_i, such that $\xi_{i,1} < \xi_{i,2} < \xi_{i,3} < \cdots$.

All these Poisson processes are defined on the probability space $(\Omega, \mathcal{F}, \mathbf{P})$ and are supposed to be mutually independent. We consider the points of all these processes as marks on $[0, \infty)$, among which we distinguish two kinds: the *cutpoints* and the *joinpoints*. The cutpoints split $[0, \infty)$ into segments, while joinpoints mark the place where to re-attach these segments. Call each point u_j a 0-cutpoint, and say that v_j is the corresponding joinpoint. Call each $\xi_{i,j}$ with $\theta_i > 0$ and $j \geq 2$ an i-cutpoint, and say that $\xi_{i,1}$ is the corresponding joinpoint. Note that for each $i \geq 1$, the mean number of i-cutpoints in the interval $[0, M]$ is $\theta_i M - 1 + e^{-\theta_i M} \leq \theta_i^2 M^2/2$, for any $M > 0$. This entails that the set of cutpoints is almost surely finite on each compact set of $[0, \infty)$, by the hypotheses on $\boldsymbol{\theta}$. In particular, we can arrange the cutpoints in increasing order as $0 < \eta_1 < \eta_2 < \eta_3 < \cdots$. We write η_k^* for the joinpoint associated to the k-th cutpoint η_k.

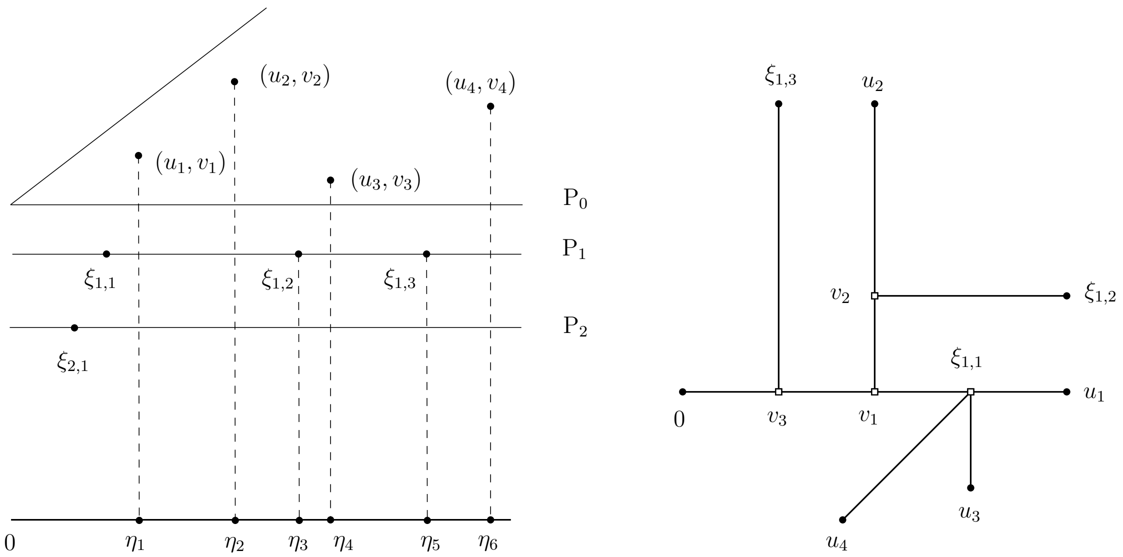

Figure 1.2 – On the left, an example of cutpoints: each η_k is either u_j or $\xi_{i,j}$ with $j \geq 2$. On the right, an example of R_6 constructed from these cutpoints. The branches attached are $[0, u_1]$, $(v_1, u_2]$, $(v_2, \xi_{1,2}]$, $(\xi_{1,1}, u_3]$, $(v_3, \xi_{1,3}]$, $(\xi_{1,1}, u_4]$.

The real tree is built by starting with the branch $[0, \eta_1]$ and inductively for $k \geq 1$, attaching the branch $(\eta_k, \eta_{k+1}]$ to the joinpoint η_k^* corresponding to the cutpoint η_k. Note that with probability 1, $\eta_k^* < \eta_k$, thus the above grafting operation is well-defined. Let R_k be the real tree obtained after attaching $(\eta_{k-1}, \eta_k]$. It is easy to see that each R_k is a real tree with the leaf set $\{0, \eta_1, \eta_2, \cdots, \eta_k\}$. See Figure 1.2 for an illustration of R_6. Furthermore, $(R_k, k \geq 1)$ forms an increasing sequence of metric spaces. Let $(\mathcal{T}, d)$ be the completion of $\cup_{k\geq 1} R_k$. Then $(\mathcal{T}, d)$ is a real tree (see Evans [57, Lemma 4.22]).

It is convenient to think of $\mathcal{T}$ as a rooted tree. We set the root of $\mathcal{T}$ at the point 0. The following properties of $\mathcal{T}$ are straightforward from the construction.

(i) The skeleton of $\mathcal{T}$ is $\cup_{k\geq 1}(\eta_k, \eta_{k+1})$.

(ii) The Lebesgue measure on $[0, \infty)$ induces a σ-finite *length measure* ℓ on $\mathcal{T}$, which assigns measure 0 to $\mathcal{T} \setminus \mathrm{Sk}(\mathcal{T})$, such that $\ell([\![x, y]\!]) = d(x, y)$ for any pair of points $x, y \in \mathcal{T}$.

(iii) The branch points of $\mathcal{T}$ correspond to the joinpoints, that is, $\mathrm{Br}(\mathcal{T}) = \{v_j : j \geq 1\} \cup \{\xi_{i,1} : \theta_i > 0, i \geq 1\}$. Each v_j has degree 3 as $u_j, j \geq 1$ are distinct almost surely, and each $\xi_{i,1}$ has infinite degree as there are infinitely many i-cutpoints.

The ICRT $\mathcal{T}$ also carries another measure, namely the *mass measure*, which is important to our study. Its definition relies on Aldous' general theory of continuum random tree (CRT) [10]. Indeed, it follows from properties of the Poisson point processes $\mathrm{P}_i, i \geq 0$, that the family $\{d(0, \eta_k), k \geq 1\}$ of root-to-leaf distances is exchangeable. Moreover, $\mathcal{T}$ satisfies the *leaf-tight* property, which amounts to say that

$$\inf_{k\geq 1} d(0, \eta_k) = 0, \qquad \text{almost surely.}$$

Actually this is guaranteed by the hypothesis that either $\theta_0 > 0$ or $\sum_{i\geq 1} \theta_i = \infty$. Then according to [10, Theorem 3], for almost every realization of $\mathcal{T}$, the empirical measure $\mu_k := \frac{1}{k}\sum_{i=1}^{k} \delta_{\eta_i}$ converges weakly as $k \to \infty$ to some probability measure μ, called the *mass measure* of $\mathcal{T}$, which is diffuse and concentrated on the leaf set. Moreover, let $(V_k, k \in \mathbb{N})$ be a sequence of independent points sampled according to μ. Then for each $k \geq 1$, the k-leafed spanning tree $\mathrm{Span}(\mathcal{T}; V_1, V_2, \cdots, V_k) := \cup_{1\leq i\leq k}[\![0, V_i]\!]$ has the same distribution as R_k. This allows us to determine the distributions of the spanning trees, as the distribution of R_k is not difficult to deduce from the construction above. We refer the reader interested in explicit formulas to Aldous and Pitman [12]. In a formal way, the equivalence class of $(\mathcal{T}, d, \mu)$ can be seen as a random variable taking values in $\mathbb{T}_w$, the space of measured metric spaces. We say the distribution of this random variable is the distribution of an *ICRT of parameter* $\boldsymbol{\theta}$.

In the case where $\boldsymbol{\theta} = (1, 0, 0, \cdots)$, the construction of $\mathcal{T}$ coincides with Algorithm 3 of Aldous [8], that is, $\mathcal{T}$ is the Brownian CRT. This is the only case where the degrees of the branch points are all finite. If $\sum_{i\geq 1} \theta_i < \infty$ (thus $\theta_0 > 0$), $\mathcal{T}$ is shown to be almost surely compact by Aldous, Miermont, and Pitman [15]. On the other hand, if $\sum_{i\geq 1} \theta_i = \infty$, the behavior of $\mathcal{T}$ can be rather wild. In this case, some heuristic arguments are proposed about a criterion for the compactness of $\mathcal{T}$ in [15]. But a mathematical justification is still missing.

ICRTs as scaling limits of birthday trees. Let $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ and let $\boldsymbol{p}_n = (p_{n1}, p_{n2}, \cdots, p_{nn})$ be a probability measure on $[n]$ for each $n \in \mathbb{N}$. We suppose further that $p_{n1} \geq p_{n2} \geq \cdots \geq p_{nn} > 0$. Let T^n be the corresponding $\boldsymbol{p}_n$-tree whose distribution is given by (1.40). Let $\sigma_n > 0$ be the number defined by $\sigma_n^2 = \sum_{i=1}^{n} p_{ni}^2$. Suppose that

$$\lim_{n\to\infty} \sigma_n = 0, \qquad \text{and} \qquad \lim_{n\to\infty} \frac{p_{ni}}{\sigma_n} = \theta_i, \qquad \text{for every } i \geq 1. \tag{1.41}$$

Let d_{T^n} be the graph distance on T^n, that is, the distance between two vertices is the number of edges on the path connecting them in T_n. Denote by $\sigma_n T^n$ the rescaled metric space $([n], \sigma_n d_{T^n})$. Camarri and Pitman [41] have shown that

$$(\sigma_n T^n, \boldsymbol{p}_n) \underset{d,\mathrm{GP}}{\overset{n\to\infty}{\longrightarrow}} (\mathcal{T}, \mu), \tag{1.42}$$

where $\to_{d,\mathrm{GP}}$ denotes the convergence in distribution with respect to Gromov–Prokhorov topology.

Remark 1. *For each $\boldsymbol{\theta} \in \boldsymbol{\Theta}$, it is not difficult to find a sequence $(\boldsymbol{p}_n, n \geq 1)$ satisfying condition (1.41) (see Aldous and Pitman [13, Lemma 4]). On the other hand, it is also shown in [41] that (1.41) is necessary to obtain a non-trivial scaling limit. Therefore, the interesting weak limits of birthday trees coincide with ICRTs.*

Remark 2. *If we take $\mathbf{p}_n$ to be the uniform measure on $[n]$ (that is, T^n is the Cayley tree with n vertices), then the limit tree is the Brownian CRT and (1.42) reduces to*

$$\frac{1}{\sqrt{n}} T^n \underset{d,\mathrm{GP}}{\overset{n\to\infty}{\longrightarrow}} \mathcal{T}^{br}. \tag{1.43}$$

This is first shown in Aldous [8] and is a special case of the convergence of the conditioned Galton–Watson in (1.39). Indeed, if we take a uniform labeling on the nodes of a Galton–Watson tree conditioned to have exactly n nodes whose offspring law is the Poisson distribution of mean 1, then we obtain in this way a Cayley tree with n vertices.

A conjecture of Aldous, Miermont and Pitman. It is conjectured in [15] that Lévy trees are mixures of ICRTs. This conjecture is motivated by the following construction of the height process of an ICRT. For $\boldsymbol{\theta} \in \boldsymbol{\Theta}$, consider the following "bridge" process with exchangeable increments:

$$Z_s^{br} = \theta_0 B_s^{br} + \sum_{i=1}^{\infty} \theta_i (\mathbf{1}_{\{U_i \leq s\}} - s), \quad 0 \leq s \leq 1,$$

where B^{br} is the Brownian bridge which returns to 0 at time 1, and $(U_i, i \geq 1)$ is a sequence of independent variables uniformly distributed on $(0, 1)$. Note that the jumps of Z^{br} have magnitude $\theta_i, i \geq 1$. Use the Vervaat transform [99], which relocates the space-time origin to the location of the infimum, to define an excursion-type process $Z = (Z_s, 0 \leq s \leq 1)$. If m denotes the Lebesgue measure on $\mathbb{R}$, let $Y = (Y_s, 0 \leq s \leq 1)$ be a continuous process defined by

$$Y_s = m\Big(\Big\{ \inf_{r\leq u\leq s} Z_u : r \leq s \Big\}\Big), \quad 0 \leq s \leq 1. \tag{1.44}$$

Then Aldous, Miermont, and Pitman [15] show that if $\boldsymbol{\theta}$ is such that $\sum_{i\geq 1} \theta_i < \infty$, then the height process of the ICRT of parameter $\boldsymbol{\theta}$ is distributed as $\frac{2}{\theta_0^2} Y$. Actually, (1.44) is an analog of a special case of (1.18): when the branching mechanism Ψ has a Brownian component, that is, $\beta > 0$ in (1.11), then (1.18) reduces to

$$H_s = \frac{1}{\beta} m\Big(\Big\{ \inf_{r\leq u\leq s} X_u : r \leq s \Big\}\Big), \quad s \geq 0, \text{ a.s.}$$

Furthermore, Kallenberg [74] has shown that a Lévy bridge process is a mixture of the extremal bridges such as Z^{br}, where the mixing measure is the distribution of the jumps in the Lévy bridge. However, the above construction of Y only works for those $\boldsymbol{\theta}$ with $\sum_{i\geq 1} \theta_i < \infty$. And it is not clear that the Vervaat transform of a general Lévy bridge would yield a "normalized Lévy excursion", though Chaumont [44] has proved that this is the case for a stable bridge process.

1.3 Main contributions of the thesis

1.3.1 Height and diameter of Brownian trees

For any integer $n \geq 1$, we denote by $\mathcal{T}_n$ a uniformly distributed random rooted labelled tree with n vertices, as defined in Section 1.2.1, and we denote by D_n its diameter with respect to the graph distance. By computations on generating functions, Szekeres [98] proved that

$$n^{-\frac{1}{2}} D_n \xrightarrow{d} \Delta \,, \tag{1.45}$$

where Δ is a random variable whose probability density f_Δ is given by

$$f_\Delta(y) = \frac{\sqrt{2\pi}}{3} \sum_{n\geq 1} \left(\frac{64}{y^4}(4b_{n,y}^4 - 36b_{n,y}^3 + 75b_{n,y}^2 - 30b_{n,y}) + \frac{16}{y^2}(2b_{n,y}^3 - 5b_{n,y}^2) \right) e^{-b_{n,y}}, \tag{1.46}$$

where $b_{n,y} := 8(\pi n/y)^2$, for all $y \in (0, \infty)$ and for all integers $n \geq 1$. This result is implicitly written in Szekeres [98] p. 395 formula (12). See also Broutin and Flajolet [38] for a similar result for binary trees. On the other hand, recall from (1.43) that $\mathcal{T}_n$, whose graph distance is rescaled by a factor $n^{-\frac{1}{2}}$, converges in distribution to the Continuum Random Tree (also called Brownian tree) that we denote by $\mathcal{T}^{br}$. From this, Aldous has deduced that Δ has the same distribution as the diameter of $\mathcal{T}^{br}$: see [9], Section 3.4, (though formula (41) there is not accurate). As proved by Aldous [10] and by Le Gall [80], the Brownian tree is coded by the normalized Brownian excursion of length 1 (see below for more details). Then, the question was raised by Aldous [9] that whether we can establish (1.46) directly from computations on the Brownian normalized excursion. In Chapter 2, we present a solution to this question: we compute the Laplace transform of law of the diameter of the Brownian tree directly from the normalized Brownian excursion and we also provide a formula for the joint law of the total height and of the diameter of the Brownian tree, which appears to be new.

Let us state more precisely our results. Recall from Section 1.1.7 the notion of stable trees conditioned on the total mass. Here, we take

$$\forall \lambda \in \mathbb{R}_+, \quad \Psi(\lambda) = \lambda^2.$$

In other words, let $X = (X_t)_{t\geq 0}$ be the underlying Lévy process whose Laplace exponent is $\Psi(\lambda) = \lambda^2$; then $(\frac{1}{\sqrt{2}} X_t)_{t\geq 0}$ is distributed as a linear standard Brownian motion such that $\mathbf{P}(X_0 = 0) = 1$. Recall the normalized excursion measure

$$\mathbf{N}_{\mathrm{nr}} = \mathbf{N}(\,\cdot\,|\,\zeta = 1) \tag{1.47}$$

as defined in (1.37). Recall that the canonical process on $\mathbb{C}(\mathbb{R}_+, \mathbb{R}_+)$ is denote by H.

Remark 3. *The positive Ito standard excursion measure $\mathbf{N}^+_{\mathrm{Ito}}$, as defined for instance in Revuz & Yor [97] Chapter XII Theorem 4.2, is derived from $\mathbf{N}$ by the following scaling relations:*

$$\mathbf{N}^+_{\mathrm{Ito}} \textit{ is the law of } \tfrac{1}{\sqrt{2}} H \textit{ under } \tfrac{1}{\sqrt{2}} \mathbf{N} \textit{ and thus, } \mathbf{N}^+_{\mathrm{Ito}}(\,\cdot\,|\,\zeta = 1) \textit{ is the law of } \tfrac{1}{\sqrt{2}} H \textit{ under } \mathbf{N}_{\mathrm{nr}}.$$

Consequently, the law $\mathbf{N}_{\mathrm{nr}}$ is not the standard normalized Brownian excursion. However, we shall call $\mathbf{N}_{\mathrm{nr}}$ the normalized Brownian excursion. $\square$

The *total height* and the *diameter* of $\mathcal{T}$ are next given by

$$\Gamma = \max_{\sigma \in \mathcal{T}} d(\rho, \sigma) = \max_{t\geq 0} H_t \quad \text{and} \quad D = \max_{\sigma, \sigma' \in \mathcal{T}} d(\sigma, \sigma') = \max_{s,t\geq 0} \big(H_t + H_s - 2 \inf_{r \in [s\wedge t, s\vee t]} H_r \big). \tag{1.48}$$

Recall that $\mathbf{m}$ stands for the mass measure on $\mathcal{T}$: it is the pushforward of the Lebesgue measure on $[0, \zeta]$ via the canonical projection $p : [0, \zeta] \to \mathcal{T}$. Recall that $\mathbf{m}(\mathcal{T}) = \zeta$.

We call *Brownian tree* the random rooted compact real tree $(\mathcal{T}, d, \rho)$ coded by H under the Brownian normalized excursion law $\mathbf{N}_{\mathrm{nr}}$. Recall from (1.28) that $\mathrm{Lf}(\mathcal{T})$ stands for the set of leaves of $\mathcal{T}$. By (1.38), we easily derive from (1.32) and (1.33) that

$$\mathbf{N}_{\mathrm{nr}}\text{-a.s.} \quad \forall \sigma \in \mathcal{T}, \quad \deg(\sigma, \mathcal{T}) \in \{1, 2, 3\}, \quad \mathbf{m} \text{ is diffuse} \quad \text{and} \quad \mathbf{m}\big(\mathcal{T} \backslash \mathrm{Lf}(\mathcal{T})\big) = 0 \,. \tag{1.49}$$

The choice of the normalizing constant $\sqrt{2}$ for the underlying Brownian motion X is explained by the following: let T^*_n be uniformly distributed on the set of rooted *planar* trees with n vertices: namely T^*_n is distributed as a Galton-Watson tree whose offspring distribution μ is geometric with mean 1 conditioned to have n vertices. Note that the variance of μ is $\sigma^2 := 2$. Thus, the convergence (1.39) and Remark 3 imply that $(n^{-\frac{1}{2}} C_{2nt}(T^*_n))_{t \in [0,1]}$ converges in law towards H under $\mathbf{N}_{\mathrm{nr}}$: see for instance Le Gall [82] Th. 1.17. Thus,

$$n^{-\frac{1}{2}} D^*_n \quad \xrightarrow{\text{(law)}} \quad D \quad \text{under} \quad \mathbf{N}_{\mathrm{nr}},$$

where D^*_n stands for the diameter of T^*_n and where D is the diameter of the Brownian tree given by (1.48).

Remark 4. *In the first paragraph of the introduction, we introduce the random tree $\mathcal{T}_n$ that is uniformly distributed on the set of rooted labeled trees with n vertices. The law of $\mathcal{T}_n$ is therefore distinct from that of T^*_n (that is uniformly distributed on the set of rooted ordered trees with n vertices). Aldous [10] has proved that the trees $\mathcal{T}_n$, whose graph distance is rescaled by a factor $n^{-\frac{1}{2}}$, converges to the tree coded by $\sqrt{2}H$, under $\mathbf{N}_{nr}$. Thus,*

$$\Delta \quad \overset{\text{(law)}}{=} \quad \sqrt{2} D \quad \textit{under} \quad \mathbf{N}_{nr} \,. \tag{1.50}$$

See Remark 6 below. □

We first prove the following result that characterizes the joint law of the height process and of the diameter of the Brownian tree.

Theorem 3 (Theorem 2.1)**.** *Recall from (1.47) the definition of the law $\mathbf{N}_{\mathrm{nr}}$ of the normalized Brownian excursion and recall from (1.48) the definition of Γ and of D. We then set*

$$\forall \lambda, y, z \in (0, \infty), \quad \mathrm{L}_\lambda(y, z) := \frac{1}{2\sqrt{\pi}} \int_0^\infty e^{-\lambda r} r^{-3/2}\, \mathbf{N}_{\mathrm{nr}}\big(r^{\frac{1}{2}} D > 2y\,;\, r^{\frac{1}{2}}\Gamma > z\big)\, dr \,. \tag{1.51}$$

Note that

$$\forall \lambda, y, z \in (0, \infty), \quad \mathrm{L}_1(y, z) = \lambda^{-\frac{1}{2}} \mathrm{L}_\lambda\big(\lambda^{-\frac{1}{2}} y\,,\, \lambda^{-\frac{1}{2}} z\big) \,. \tag{1.52}$$

Then,

$$\mathrm{L}_1(y, z) = \coth(y \vee z) - 1 - \tfrac{1}{4} \mathbf{1}_{\{z \le 2y\}} \frac{\sinh(2q) - 2q}{\sinh^4(y)} \,, \tag{1.53}$$

where $q = y \wedge (2y - z)$. In particular, this implies that

$$\forall \lambda, z \in (0, \infty), \quad \mathrm{L}_\lambda(0, z) = \sqrt{\lambda} \coth(z\sqrt{\lambda}) - \sqrt{\lambda} \,. \tag{1.54}$$

and

$$\forall \lambda, y \in (0, \infty), \quad \mathrm{L}_\lambda(y, 0) = \sqrt{\lambda} \coth(y\sqrt{\lambda}) - \sqrt{\lambda} - \sqrt{\lambda}\, \frac{\sinh(2y\sqrt{\lambda}) - 2y\sqrt{\lambda}}{4 \sinh^4(y\sqrt{\lambda})} \,, \tag{1.55}$$

From this theorem we deduce the following explicit laws.

Corollary 4 (Corollary 2.2)**.** *For all $y, z \in (0,\infty)$, we set*

$$\rho = z \vee \tfrac{y}{2} \quad \textit{and} \quad \delta = \left(\tfrac{2(y-z)}{y} \vee 0\right) \wedge 1 \, . \tag{1.56}$$

Then we have

$$\mathbf{N}_{\mathrm{nr}}\big(D > y\,;\, \Gamma > z\big) = 2\sum_{n\geq 1} \big(2n^2\rho^2 - 1\big)\, e^{-n^2\rho^2} + \tag{1.57}$$

$$\frac{1}{6}\sum_{n\geq 2} n(n^2-1)\Big[\big[(n+\delta)^2y^2-2\big]e^{-\frac{1}{4}(n+\delta)^2y^2} - \big[(n-\delta)^2y^2-2\big]e^{-\frac{1}{4}(n-\delta)^2y^2} + \delta y(n^3y^3-6ny)e^{-\frac{1}{4}n^2y^2}\Big]$$

and

$$\begin{aligned}\mathbf{N}_{\mathrm{nr}}\big(D \leq y\,;\, \Gamma \leq z\big) = {} & \frac{4\pi^{5/2}}{\rho^3}\sum_{n\geq 1} n^2 e^{-n^2\pi^2/\rho^2} - \\ & \frac{32\pi^{3/2}}{3}\sum_{n\geq 1} n\sin(2\pi n\delta)\Big(\frac{2}{y^5}(2a_{n,y}^2 - 9a_{n,y} + 6) - \frac{3\delta^2-1}{y^3}(a_{n,y}-1)\Big)e^{-a_{n,y}} + \\ & \frac{16\pi^{1/2}}{3}\sum_{n\geq 1} \delta\cos(2\pi n\delta)\Big(\frac{1}{y^3}(6a_{n,y}^2 - 15a_{n,y} + 3) - \frac{\delta^2-1}{2y}a_{n,y}\Big)e^{-a_{n,y}} + \\ & \frac{16\pi^{1/2}}{3}\sum_{n\geq 1} \delta\Big(\frac{1}{y^3}(4a_{n,y}^3 - 24a_{n,y}^2 + 27a_{n,y} - 3) + \frac{1}{2y}(2a_{n,y}^2 - 3a_{n,y})\Big)e^{-a_{n,y}},\end{aligned} \tag{1.58}$$

where we set $a_{n,y} = 4(\pi n/y)^2$ for all $y \in (0,\infty)$ and for all $n \geq 1$ to simplify notation. In particular, (1.57) *implies*

$$\mathbf{N}_{\mathrm{nr}}\big(\Gamma > y\big) = 2\sum_{n\geq 1}\big(2n^2y^2 - 1\big)\, e^{-n^2y^2}, \tag{1.59}$$

and

$$\mathbf{N}_{\mathrm{nr}}\big(D > y\big) = \sum_{n\geq 2}(n^2-1)\big(\tfrac{1}{6}n^4y^4 - 2n^2y^2 + 2\big)e^{-n^2y^2/4}. \tag{1.60}$$

On the other hand, (1.58) *implies*

$$\mathbf{N}_{\mathrm{nr}}\big(\Gamma \leq y\big) = \frac{4\pi^{5/2}}{y^3}\sum_{n\geq 1} n^2 e^{-n^2\pi^2/y^2}, \tag{1.61}$$

and

$$\mathbf{N}_{\mathrm{nr}}\big(D \leq y\big) = \frac{\sqrt{\pi}}{3}\sum_{n\geq 1}\Big(\frac{8}{y^3}(24a_{n,y} - 36a_{n,y}^2 + 8a_{n,y}^3) + \frac{16}{y}a_{n,y}^2\Big)e^{-a_{n,y}}\, . \tag{1.62}$$

Thus the law of D under $\mathbf{N}_{\mathrm{nr}}$ has the following density:

$$\begin{aligned} f_D(y) &= \frac{1}{12}\sum_{n\geq 1}\big(n^8y^5 - n^6y^3(20+y^2) + 20n^4y(3+y^2) - 60n^2y\big)e^{-n^2y^2/4} &(1.63)\\ &= \frac{2\sqrt{\pi}}{3}\sum_{n\geq 1}\Big(\frac{16}{y^4}(4a_{n,y}^4 - 36a_{n,y}^3 + 75a_{n,y}^2 - 30a_{n,y}) + \frac{8}{y^2}(2a_{n,y}^3 - 5a_{n,y}^2)\Big)e^{-a_{n,y}}\, . &(1.64)\end{aligned}$$

Remark 5. *We derive* (1.58) *from* (1.57) *using the following identity on the theta function due to Jacobi (1828), which is a consequence of Poisson summation formula:*

$$\forall t \in (0,\infty), \forall x, y \in \mathbb{C}, \quad \sum_{n\in\mathbb{Z}} e^{-(x+n)^2t - 2\pi i n y} = e^{2\pi i x y}\left(\frac{\pi}{t}\right)^{\frac{1}{2}}\sum_{n\in\mathbb{Z}} e^{-\frac{\pi^2(y+n)^2}{t} + 2\pi i n x}\, . \tag{1.65}$$

See for instance Weil [101], Chapter VII, Equation (12). Not surprisingly, (1.65) *can also be used to derive (1.61) from (1.59), to derive (1.62) from (1.60), or to derive (1.64) from (1.63).* □

Remark 6. *We obtain ([1.63](#)) (resp. ([1.64](#))) by differentiating ([1.60](#)) (resp. ([1.62](#))). By ([1.50](#)), we have*

$$\forall y \in (0,\infty), \quad f_\Delta(y) = \frac{1}{\sqrt{2}} f_D\Big(\frac{y}{\sqrt{2}}\Big) ,$$

which immediately entails (1.46) from (1.64), since $a_{n,y/\sqrt{2}} = 8(\pi n/y)^2 = b_{n,y}$. □

1.3.2 Decomposition of Lévy trees along their diameter

In the article [54] (that is written with Duquesne and that corresponds to Chapter 3), we compute the law of the diameter for general Lévy trees (see Theorem 5). We also prove that the diameter of Lévy trees is realized by a unique pair of points. The geodesic path joining these two extremal points is therefore unique. In Theorem 6, we describe the coding function (the height process) of the Lévy trees rerooted at the midpoint of their diameter, which plays the role of an intrinsic root. The proof of Theorem 6 that provides a decomposition of Lévy trees according to their diameter specifically relies on the invariance of Lévy trees by uniform rerooting, as proved by Duquesne and Le Gall [53], and on the decomposition of Lévy trees according to their height, as proved by Abraham and Delmas [3] (this decomposition generalizes Williams' decomposition of the Brownian excursion). Roughly speaking, Theorem 6 asserts that a Lévy tree that is conditioned to have diameter r and that is rooted at its midpoint is obtained by glueing at their root two size-biased independent Lévy trees conditioned to have height $r/2$ and then by rerooting uniformly the resulting tree; Theorem 6 also explains the distribution of the trees grafted on the diameter. As an application of this theorem, we characterize the joint law of the height and the diameter of stable trees conditioned by their total mass (see Proposition 7) and by providing an asymptotic expansion of the law of the height (Theorem 9) and of the law of the diameter (Theorem 11). These two asymptotic expansions generalize the identities of Szekeres in the Brownian case which involves theta functions (see (1.59) and (1.60)).

Before stating precisely these results, we need to introduce definitions and notations.

Re-rooting trees. Several statements involve a re-rooting procedure at the level of the coding functions that is recalled here from Duquesne and Le Gall [52], Lemma 2.2 (see also Duquesne and Le Gall [53]). Let H be a coding function as defined in Section 1.1.6 and recall that $\zeta_H \in (0,\infty)$. For any $t \in \mathbb{R}_+$, denote by $\bar{t}$ the unique element of $[0,\zeta_H)$ such that $t - \bar{t}$ is an integer multiple of ζ_H. Then for all $t_0 \in \mathbb{R}_+$, we set

$$\forall t \in [0,\zeta_H], \quad H_t^{[t_0]} = d_H\big(\overline{t_0}, \overline{t+t_0}\big) \quad \text{and} \quad \forall t \geq \zeta_H, \quad H_t^{[t_0]} = 0 . \tag{1.66}$$

Then observe that $\zeta_H = \zeta_{H^{[t_0]}}$ and that

$$\forall t,t' \in [0,\zeta_H], \quad d_{H^{[t_0]}}(t,t') = d_H\big(\overline{t+t_0}, \overline{t'+t_0}\big) . \tag{1.67}$$

Lemma 2.2 [52] asserts that there exists a unique isometry $\phi : \mathcal{T}_{H^{[t_0]}} \to \mathcal{T}_H$ such that $\phi(p_{H^{[t_0]}}(t)) = p_H\big(\overline{t+t_0}\big)$ for all $t \in [0,\zeta_H]$. This allows to *identify canonically* $\mathcal{T}_{H^{[t_0]}}$ *with the tree* $\mathcal{T}_H$ *re-rooted at* $p_H(t_0)$:

$$\big(\mathcal{T}_{H^{[t_0]}}, d_{H^{[t_0]}}, \rho_{H^{[t_0]}}\big) \equiv \big(\mathcal{T}_H, d_H, p_H(t_0)\big) . \tag{1.68}$$

Note that up to this identification, $\mathbf{m}_{H^{[t_0]}}$ is the same as $\mathbf{m}_H$. Roughly speaking, the linear order $\leq_{H^{[t_0]}}$ is obtained from $\leq_H$ by a cyclic shift after $p_H(t_0)$.

Spinal decomposition. The law of the Lévy tree conditioned by its diameter that is discussed below is described as a Poisson decomposition of the trees grafted along the diameter. To explain that kind of decomposition in terms of the coding function of the tree, we introduce the following definitions and notations.

Let $h\in\mathbb{C}(\mathbb{R}_+,\mathbb{R}_+)$ have compact support. Note that $h(0)>0$ possibly. We first define the excursions of h above its infimum as follows. For any $a\in[0,h(0)]$, we first set

$$\ell_a(h) := \inf\big\{t\in\mathbb{R}_+ : h(t)=h(0)-a\big\} \quad \text{and} \quad r_a(h) := \zeta_h \wedge \inf\big\{t\in(0,\infty) : h(0)-a > h(t)\big\},$$

with the convention that $\inf\emptyset=\infty$, so that $r_{h(0)}(h)=\zeta_h$. We then set

$$\forall s\in\mathbb{R}_+, \quad \mathcal{E}_s(h,a) := h\big((\ell_a(h)+s)\wedge r_a(h)\big) - h(0) + a \, .$$

See Figure 1.3. Note that $\mathcal{E}(h,a)$ is a nonnegative continuous function with compact support such that $\mathcal{E}_0(h,a)=0$. Moreover, if $\ell_a(h)=r_a(h)$, then $\mathcal{E}(h,a)=\mathbf{0}$, the *null function*.

Let H be a coding function as defined above. Let $t\in\mathbb{R}_+$. We next set

$$\forall s\in\mathbb{R}_+, \quad H^-_s = H_{(t-s)_+} \quad \text{and} \quad H^+_s = H_{t+s} \, .$$

Note that $H^-_0=H^+_0=H_t$. To simplify notation we also set

$$\forall a\in[0,H_t], \quad \overleftarrow{H}^a := \mathcal{E}(H^-,a) \quad \text{and} \quad \overrightarrow{H}^a := \mathcal{E}(H^+,a)$$

and

$$\mathcal{J}_{0,t} := \big\{a\in[0,H_t] : \text{either } \ell_a(H^-)<r_a(H^-) \text{ or } \ell_a(H^+)<r_a(H^+)\big\}$$

that is countable. We then define the following point measure on $[0,H_t]\times\mathbb{C}(\mathbb{R}_+,\mathbb{R}_+)^2$:

$$\mathcal{M}_{0,t}(H) = \sum_{a\in\mathcal{J}_{0,t}} \delta_{(a,\overleftarrow{H}^a,\overrightarrow{H}^a)} \, , \tag{1.69}$$

with the convention that $\mathcal{M}_{0,t}(H)=0$ if $\mathcal{J}_{0,t}=\emptyset$.

For all $t_1\geq t_0\geq 0$, we also set

$$\mathcal{M}_{t_0,t_1}(H) := \mathcal{M}_{0,t_1-t_0}\big(H^{[t_0]}\big) =: \sum_{a\in\mathcal{J}_{t_0,t_1}} \delta_{(a,\overleftarrow{H}^a,\overrightarrow{H}^a)} \, . \tag{1.70}$$

This point measure on $[0,d_H(t_0,t_1)]\times\mathbb{C}(\mathbb{R}_+,\mathbb{R}_+)^2$ is the *spinal decomposition of H between t_0 and t_1*.

Let us interpret this decomposition in terms of the tree $\mathcal{T}_H$ (see Figure 1.3): set $\gamma_0=p_H(t_0)$ and $\gamma_1=p_H(t_1)$; to simplify, we assume that γ_0 and γ_1 are leaves. Recall that $[\![\gamma_0,\gamma_1]\!]$ is the geodesic path joining γ_0 and γ_1; then $\mathcal{J}_{t_0,t_1}=\{d(\sigma,\gamma_1);\sigma\in\mathtt{Br}(\mathcal{T}_H)\cap[\![\gamma_0,\gamma_1]\!]\}$. For any positive $a\in\mathcal{J}_{t_0,t_1}$, there exists $\sigma\in\mathtt{Br}(\mathcal{T}_H)\cap[\![\gamma_0,\gamma_1]\!]$ such that the following holds true.

- $\overrightarrow{\mathcal{T}}_a := \{\sigma\}\cup\big\{\sigma'\in\mathcal{T}_H : \gamma_0<_H\sigma'<_H\gamma_1 \text{ and } [\![\gamma_0,\sigma]\!]=[\![\gamma_0,\sigma']\!]\cap[\![\gamma_0,\gamma_1]\!]\big\}$ is the tree grafted at σ on the right hand side of $[\![\gamma_0,\gamma_1]\!]$ and the tree $(\overrightarrow{\mathcal{T}}_a,d,\sigma)$ is coded by $\overrightarrow{H}^a$.
- $\overleftarrow{\mathcal{T}}_a := \{\sigma\}\cup\big\{\sigma'\in\mathcal{T}_H : \text{either } \sigma'<_H\gamma_0 \text{ or } \gamma_1<_H\sigma' \text{ and } [\![\gamma_0,\sigma]\!]=[\![\gamma_0,\sigma']\!]\cap[\![\gamma_0,\gamma_1]\!]\big\}$ is the tree grafted at σ on the left hand side of $[\![\gamma_0,\gamma_1]\!]$ and the tree $(\overleftarrow{\mathcal{T}}_a,d,\sigma)$ is coded by $\overleftarrow{H}^a$.

Diameter decomposition. Let Ψ be a branching mechanism of the form (1.11) that satisfies (1.14). Recall that X stands for a spectrally positive Lévy process defined on $(\Omega,\mathcal{F},\mathbf{P})$ starting from 0 and whose Laplace exponent is Ψ: see (1.17). Recall from (1.18) the definition of the Ψ-height process H under $\mathbf{P}$ and under its excursion measure $\mathbf{N}$. Recall that the tree $\mathcal{T}$ coded by H under $\mathbf{N}$ is the Ψ-Lévy tree. One checks that the total height is $\mathbf{N}$-a.s. realized at a unique time (see Duquesne and Le Gall [52] and also Abraham and Delmas [3]). Namely,

$$\mathbf{N}\text{-a.e. there exists a unique } \tau\in[0,\zeta] \text{ such that } H_\tau=\Gamma \, . \tag{1.71}$$

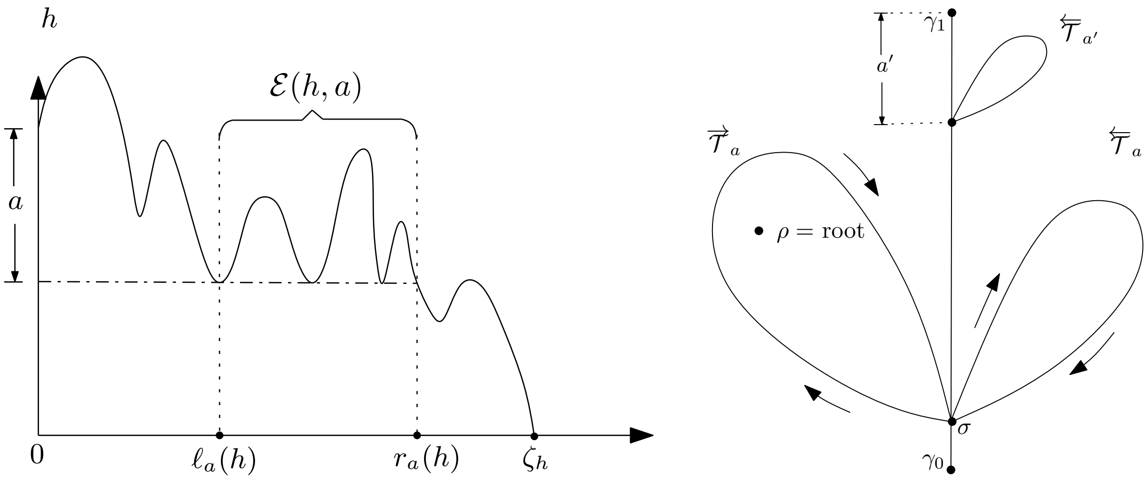

Figure 1.3 – *The figure on the left hand side illustrates the definition of $\mathcal{E}(h,a)$. The figure on the right hand side represents the spinal decomposition of H at times t_0 and t_1 in terms of the tree $\mathcal{T}$ coded by H.*

Moreover, the distribution of the total height Γ under $\mathbf{N}$ is characterized as follows:

$$\forall t \in (0,\infty), \quad v(t) := \mathbf{N}(\Gamma > t) \quad \text{satisfies} \quad \int_{v(t)}^{\infty} \frac{d\lambda}{\Psi(\lambda)} = t\,. \tag{1.72}$$

Recall from (1.16) that $v:(0,\infty)\to(0,\infty)$ is a bijective decreasing C^∞ function and (1.72) implies that on $(0,\infty)$, $\mathbf{N}(\Gamma\in dt)=\Psi(v(t))\,dt$.

For all $x\in(0,\infty)$, we set $T_x=\inf\{t\in\mathbb{R}_+ : X_t=-x\}$ that is $\mathbf{P}$-a.s. finite since X under $\mathbf{P}$ does not drift to ∞. We next introduce the following law $\mathbf{P}^x$ on $\mathbb{C}(\mathbb{R}_+,\mathbb{R}_+)$:

$$\mathbf{P}^x \text{ is the law of } (H_{t\wedge T_x})_{t\geq 0} \text{ under } \mathbf{P}, \tag{1.73}$$

The tree $\mathcal{T}_H$ under $\mathbf{P}^x(dH)$ is called the *Ψ-Lévy forest starting from a population of size x*. Then, the mass measure of $\mathcal{T}_H$ under $\mathbf{P}^x(dH)$ satisfies the following important properties:

$$\mathbf{P}^x(dH)\text{-a.s. } \mathbf{m}_H \text{ is diffuse and } \mathbf{m}_H(\mathcal{T}_H\backslash \mathtt{Lf}(\mathcal{T}_H))=0. \tag{1.74}$$

The Poisson decomposition (1.25) implies that $\sup_{t\in[0,T_x]} H_t=\max\{\Gamma(H^i); i\in\mathcal{I} : -I_{a_i}\leq x\}$ and since Γ under $\mathbf{N}$ has a density, then (1.71) and (1.72) entail that

$$\mathbf{P}^x\text{-a.s. there is a unique } \tau\in[0,\zeta] \text{ such that } H_\tau=\Gamma \quad \text{and} \quad \mathbf{P}^x(\Gamma\leq t)=e^{-xv(t)},\, t\in\mathbb{R}_+. \tag{1.75}$$

In [3], Abraham and Delmas generalize Williams' decomposition of the Brownian excursion to the excursion of the Ψ-height process: they first make sense of the conditioned law $\mathbf{N}(\,\cdot\,|\,\Gamma=r)$. Namely they prove that $\mathbf{N}(\,\cdot\,|\,\Gamma=r)$-a.s. $\Gamma=r$, that $r\mapsto\mathbf{N}(\,\cdot\,|\,\Gamma=r)$ is weakly continuous on $\mathbb{C}(\mathbb{R}_+,\mathbb{R}_+)$ and that

$$\mathbf{N}=\int_0^\infty \mathbf{N}(\Gamma\in dr)\,\mathbf{N}(\,\cdot\,|\,\Gamma=r)\,. \tag{1.76}$$

Moreover they provide a Poisson decomposition along the total height of the process: see Chapter 3, Section 3.2.2 where a more precise statement is recalled. The first two results of our article provide a similar result for the diameter D of the Ψ-Lévy tree under $\mathbf{N}$. Recall that $p:[0,\zeta]\to\mathcal{T}$ stands for the canonical projection.

Theorem 5 (Theorem 3.1)**.** *Let Ψ be a branching mechanism of the form (1.11) that satisfies (1.14). Let $\mathcal{T}$ be the Ψ-Lévy tree that is coded by the Ψ-height process H under the excursion measure $\mathbf{N}$ as defined above. Then, the following holds true $\mathbf{N}$-a.e.*

(i) There exists a unique pair $\tau_0, \tau_1 \in [0,\zeta]$ such that $\tau_0 < \tau_1$ and $D = d(\tau_0,\tau_1)$. Moreover, either $H_{\tau_0} = \Gamma$ or $H_{\tau_1} = \Gamma$. Namely, either $\tau_0 = \tau$ or $\tau_1 = \tau$, where τ is the unique time realizing the total height as defined by (1.71).

(ii) Set $\gamma_0 = p(\tau_0)$ and $\gamma_1 = p(\tau_1)$. Then γ_0 and γ_1 are leaves of $\mathcal{T}$. Let γ_{mid} be the mid-point of $[\![\gamma_0,\gamma_1]\!]$: namely, γ_{mid} is the unique point of $[\![\gamma_0,\gamma_1]\!]$ such that $d(\gamma_0,\gamma_{\mathrm{mid}}) = D/2$. Then, there are exactly two times $0 \leq \tau^-_{\mathrm{mid}} < \tau^+_{\mathrm{mid}} \leq \zeta$ such that $p(\tau^-_{\mathrm{mid}}) = p(\tau^+_{\mathrm{mid}}) = \gamma_{\mathrm{mid}}$, and γ_{mid} is a simple point of $\mathcal{T}$: namely, it is neither a branching point nor a leaf of $\mathcal{T}$.

(iii) For all $r \in (0,\infty)$, we get

$$\mathbf{N}\big(D > 2r\big) = v(r) - \Psi\big(v(r)\big)^2 \int_{v(r)}^{\infty} \frac{d\lambda}{\Psi(\lambda)^2} \, . \tag{1.77}$$

This implies that $\mathbf{N}(D \in dr) = \varphi(r)dr$ on $(0,\infty)$ where the density $\varphi : (0,\infty) \to (0,\infty)$ is given by

$$\forall r \in (0,\infty), \quad \varphi(2r) = \Psi(v(r)) - \Psi(v(r))^2\, \Psi'(v(r)) \int_{v(r)}^{\infty} \frac{d\lambda}{\Psi(\lambda)^2} \, . \tag{1.78}$$

The second main result of our paper is a Poisson decomposition of the subtrees of $\mathcal{T}$ grafted on the diameter $[\![\gamma_0,\gamma_1]\!]$. This result is stated in terms of coding functions and we first need to introduce the following notation: let $H, H' \in \mathbb{C}(\mathbb{R}_+,\mathbb{R}_+)$ be two coding functions as defined above; the *concatenation* of H and H' is the coding function denoted by $H \oplus H'$ and given by

$$\forall t \in \mathbb{R}_+, \qquad (H \oplus H')_t = H_t \quad \text{if } t \in [0,\zeta_H] \quad \text{and} \quad (H \oplus H')_t = H'_{t-\zeta_H} \quad \text{if } t \geq \zeta_H. \tag{1.79}$$

Moreover, to simplify notation we write the following:

$$\forall r \in (0,\infty), \quad \mathbf{N}^{\Gamma}_r = \mathbf{N}(\,\cdot \,|\, \Gamma = r) \, . \tag{1.80}$$

Theorem 6 (Theorem 3.2)**.** *Let Ψ be a branching mechanism of the form (1.11) that satisfies (1.14). For all $r \in (0,\infty)$, we denote by $\mathbf{Q}_r$ the law on $\mathbb{C}(\mathbb{R}_+,\mathbb{R}_+)$ of $H \oplus H'$ under $\mathbf{N}^{\Gamma}_{r/2}(dH)\mathbf{N}^{\Gamma}_{r/2}(dH')$, where $\mathbf{N}^{\Gamma}_{r/2}$ is defined by (1.80). Namely, for all measurable functions $F : \mathbb{C}(\mathbb{R}_+,\mathbb{R}_+) \to \mathbb{R}_+$,*

$$\mathbf{Q}_r\big[F(H)\big] = \iint_{\mathbb{C}(\mathbb{R}_+,\mathbb{R}_+)^2} \mathbf{N}^{\Gamma}_{r/2}(dH)\mathbf{N}^{\Gamma}_{r/2}(dH')\; F\big(H \oplus H'\big) \, . \tag{1.81}$$

Then $\mathbf{Q}_r$ satisfies the following properties.

(i) $\mathbf{Q}_r$-a.s. $D = r$ and there exists a unique pair of points $\tau_0, \tau_1 \in [0,\zeta]$ such that $D = d(\tau_0,\tau_1)$.

(ii) For all $r \in (0,\infty)$, $\mathbf{Q}_r[\,\zeta\,] = 2\mathbf{N}^{\Gamma}_{r/2}[\,\zeta\,] \in (0,\infty)$. Moreover, the application $r \mapsto \mathbf{Q}_r$ is weakly continuous and for all measurable functions $F : \mathbb{C}(\mathbb{R}_+,\mathbb{R}_+) \to \mathbb{R}_+$ and $f : \mathbb{R}_+ \to \mathbb{R}_+$,

$$\mathbf{N}\big[f(D)F(H)\big] = \int_0^{\infty} \frac{\mathbf{N}(D \in dr)}{\mathbf{Q}_r[\,\zeta\,]}\, f(r)\, \mathbf{Q}_r\Big[\int_0^{\zeta} F\big(H^{[t]}\big)\, dt\Big] \, , \tag{1.82}$$

where $H^{[t]}$ is defined by (1.66).

(iii) Recall the notation τ^-_{mid} and τ^+_{mid} from Theorem 5 (ii). Then, for all $r \in (0,\infty)$,

$$\mathbf{N}\big[F\big(H^{[\tau^-_{\mathrm{mid}}]}\big) \,\big|\, D = r\big] = \frac{1}{\mathbf{N}^{\Gamma}_{r/2}[\,\zeta\,]} \iint_{\mathbb{C}(\mathbb{R}_+,\mathbb{R}_+)^2} \mathbf{N}^{\Gamma}_{r/2}(dH)\mathbf{N}^{\Gamma}_{r/2}(dH')\; \zeta_{H'} F\big(H \oplus H'\big) \, , \tag{1.83}$$

where $\mathbf{N}(\,\cdot\,|\, D = r)$ makes sense for all $r \in (0,\infty)$ thanks to (1.82).

(iv) *Recall from (1.73) the notation* $\mathbf{P}^y$*. To simplify notation, we write for all* $y, b \in (0, \infty)$

$$\mathbf{N}_b = \mathbf{N}\big(\,\cdot \cap \{\Gamma \leq b\}\big) \quad \textit{and} \quad \mathbf{P}^y_b = \mathbf{P}^y\big(\,\cdot \cap \{\Gamma \leq b\}\big), \tag{1.84}$$

Then, under $\mathbf{Q}_r$*,* $\mathcal{M}_{\tau_0,\tau_1}(da\, d\overleftarrow{H}\, d\overrightarrow{H})$*, defined by (1.70), is a Poisson point measure on* $[0,r] \times \mathbb{C}(\mathbb{R}_+,\mathbb{R}_+)^2$ *whose intensity is*

$$\begin{aligned}\beta \mathbf{1}_{[0,r]}(a) da \, \Big(\delta_{\mathbf{0}}(d\overleftarrow{H}) \mathbf{N}_{a\wedge(r-a)}(d\overrightarrow{H}) + \mathbf{N}_{a\wedge(r-a)}(d\overleftarrow{H}) \delta_{\mathbf{0}}(d\overrightarrow{H}) \Big) \\ + \; \mathbf{1}_{[0,r]}(a) da \int_{(0,\infty)} \pi(dz) \int_0^z dx \; \mathbf{P}^x_{a\wedge(r-a)}\big(d\overleftarrow{H}\big) \, \mathbf{P}^{z-x}_{a\wedge(r-a)}\big(d\overrightarrow{H}\big),\end{aligned} \tag{1.85}$$

where β *and* π *are defined in (1.11).*

Remark 7. *As already mentioned, the previous theorem makes sense of* $\mathbf{N}\big(\,\cdot \,\big| D = r\big)$ *and for all measurable functions* $F : \mathbb{C}(\mathbb{R}_+,\mathbb{R}_+) \to \mathbb{R}_+$*, we have*

$$\forall r \in (0,\infty), \qquad \mathbf{N}\big[\, F(H) \,\big|\, D = r\big] := \mathbf{Q}_r\Big[\int_0^\zeta F\big(H^{[t]}\big)\, dt \Big] \Big/ \mathbf{Q}_r[\,\zeta\,] \, , \tag{1.86}$$

Namely, Theorem 6 (i) *entails that* $\mathbf{N}(\,\cdot\,\big|\, D = r)$*-a.s.* $D = r$*. Then (1.81) combined with the already mentioned continuity of* $r \mapsto \mathbf{N}(\,\cdot\,|\,\Gamma = r/2)$ *easily implies that* $r \mapsto \mathbf{N}(\,\cdot\,\big|\, D = r)$ *is weakly continuous on* $\mathbb{C}(\mathbb{R}_+,\mathbb{R}_+)$*. Moreover, (1.82) can be rewritten:*

$$\mathbf{N} = \int_0^\infty \mathbf{N}(D \in dr)\, \mathbf{N}(\,\cdot\,|\, D = r) \tag{1.87}$$

that is the exact analogous of (1.76). We mention that the proof of Theorem 6 relies on the decomposition (1.76) due to Abraham and Delmas [3]. $\square$

Remark 8. *It is easy to check from (1.66) that for all* t_0, t*,* $(H^{[t]})^{[t_0]} = H^{[t+t_0]}$*. Therefore, (1.82) implies that* H *under* $\mathbf{N}$ *is invariant under rerooting. Namely, for all measurable functions* $F : \mathbb{C}(\mathbb{R}_+,\mathbb{R}_+) \to \mathbb{R}_+$*,*

$$\forall t_0 \in \mathbb{R}_+, \quad \mathbf{N}\big[\mathbf{1}_{\{\zeta \geq t_0\}} F\big(H^{[t_0]}\big)\big] = \mathbf{N}\big[\mathbf{1}_{\{\zeta \geq t_0\}} F\big(H\big)\big] \, , \tag{1.88}$$

which is quite close to Proposition 2.1 in Duquesne and Le Gall [53], that is used in the proof of Theorem 6. $\square$

Remark 9. *As shown by (1.86),* $\mathbf{N}\big(\,\cdot \,\big| D = r\big)$ *is derived from* $\mathbf{Q}_r$ *by a uniform rerooting. This property suggests that the law of the compact real tree* $(\mathcal{T}, d)$ *coded by* H *under* $\mathbf{Q}_r$*, without its root, is the scaling limit of natural models of labeled unrooted trees conditioned by their diameter.* $\square$

Remark 10. *Another reason for introducing the law* $\mathbf{Q}_r$ *is the following: we deduce from (1.86) that for all measurable functions* $F : \mathbb{C}(\mathbb{R}_+,\mathbb{R}_+) \to \mathbb{R}_+$*,*

$$\mathbf{N}\big[F(H^{[\tau_0]}) \,\big|\, D = r\big] = \mathbf{Q}_r\big[\zeta F(H^{[\tau_0]})\big] \big/ \mathbf{Q}_r[\,\zeta\,] \, , \tag{1.89}$$

where τ_0 *is as in Theorem 5. By Theorem 6* (iv)*,* H *under* $\mathbf{Q}_r$ *enjoys a Poisson decomposition along its diameter. However (1.89) implies that this not the case of* H *under* $\mathbf{N}(\,\cdot\,|\, D = r)$*.* $\square$

The law of Γ and of D of stable Lévy trees conditioned by their total mass. In application of Theorem 6, we compute the law of Γ and D under $\mathbf{N}(\,\cdot\,|\,\zeta=1)$ in the cases where Ψ is a stable branching mechanism. Namely, we fix $\gamma\in(1,2]$ and

$$\Psi(\lambda)=\lambda^{\gamma},\quad \lambda\in\mathbb{R}_+\ ,$$

that is called the γ*-stable branching mechanism*. Recall from (1.37) the definition of

$$\mathbf{N}_{\mathrm{nr}}:=\mathbf{N}(\,\cdot\,|\,\zeta=1)\ .$$

We next introduce $w:(0,\infty)\to(1,\infty)$ that is the unique C^{∞} decreasing bijection that satisfies the following integral equation:

$$\forall y\in(0,\infty),\qquad \int_{w(y)}^{\infty}\frac{du}{u^{\gamma}-1}=y\ . \tag{1.90}$$

We refer to Chapter 3, Section 3.3.1 for a probabilistic interpretation of w and further properties. The following proposition characterizes the joint law of Γ and D under $\mathbf{N}_{\mathrm{nr}}$ by the mean of Laplace transforms.

Proposition 7 (Proposition 3.3)**.** *Fix $\gamma\in(1,2]$ and $\Psi(\lambda)=\lambda^{\gamma}$, $\lambda\in\mathbb{R}_+$. Recall from (1.37) the definition of the law $\mathbf{N}_{\mathrm{nr}}$ of the normalized excursion of the γ-stable height process. We then set*

$$\forall\lambda,y,z\in(0,\infty),\quad \mathrm{L}_{\lambda}(y,z):=c_{\gamma}\int_0^{\infty}e^{-\lambda r}r^{-1-\frac{1}{\gamma}}\,\mathbf{N}_{\mathrm{nr}}\big(r^{\frac{\gamma-1}{\gamma}}D>2y\,;\,r^{\frac{\gamma-1}{\gamma}}\Gamma>z\big)\,dr\ , \tag{1.91}$$

where we recall from (1.35) that $1/c_{\gamma}=\gamma\Gamma_e\big(\frac{\gamma-1}{\gamma}\big)$, Γ_e standing for Euler's Gamma function. Note that

$$\forall\lambda,y,z\in(0,\infty),\quad \mathrm{L}_1(y,z)=\lambda^{-\frac{1}{\gamma}}\mathrm{L}_{\lambda}\big(\lambda^{-\frac{\gamma-1}{\gamma}}y\,,\,\lambda^{-\frac{\gamma-1}{\gamma}}z\big)\ . \tag{1.92}$$

Recall from (1.90) the definition of w. Then,

$$\mathrm{L}_1(y,z)=w(y\vee z)-1-\tfrac{1}{\gamma}\mathbf{1}_{\{z<2y\}}\,\big(w(y)^{\gamma}-1\big)^2\left(\frac{w\big(y\wedge(2y-z)\big)}{w\big(y\wedge(2y-z)\big)^{\gamma}-1}-(\gamma-1)\big(y\wedge(2y-z)\big)\right). \tag{1.93}$$

In particular, for all $y,z\in(0,\infty)$,

$$\mathrm{L}_1(0,z)=w(z)-1\quad\textit{and}\quad \mathrm{L}_1(y,0)=w(y)-1-\tfrac{1}{\gamma}\big(w(y)^{\gamma}-1\big)\Big(w(y)-(\gamma-1)y\big(w(y)^{\gamma}-1\big)\Big). \tag{1.94}$$

The previous proposition is known in the Brownian case, where $w(y)=\coth(y)$: see Section 1.3.1. As already mentioned in Corollary 4, in the Brownian case, standard computations derived from (1.94) imply the following power expansions that hold true for all $y\in(0,\infty)$:

$$\mathbf{N}_{\mathrm{nr}}\big(\Gamma>y\big)=2\sum_{n\geq1}\big(2n^2y^2-1\big)\,e^{-n^2y^2}\ , \tag{1.95}$$

$$\mathbf{N}_{\mathrm{nr}}\big(D>y\big)=\sum_{n\geq2}(n^2-1)\big(\tfrac{1}{6}n^4y^4-2n^2y^2+2\big)e^{-n^2y^2/4}\ . \tag{1.96}$$

We next provide similar asymptotic expansions in the non-Brownian stable cases. To that end, we introduce $s_{\gamma}:(0,\infty)\to(0,\infty)$ as the continuous version of the density of the spectrally positive $\frac{\gamma-1}{\gamma}$-stable distribution; more precisely, s_{γ} is characterized by the following:

$$\forall\lambda\in(0,\infty),\qquad \int_0^{\infty}e^{-\lambda x}s_{\gamma}(x)\,dx=\exp(-\gamma\lambda^{\frac{\gamma-1}{\gamma}})\ . \tag{1.97}$$

The following asymptotic expansion of s_γ at 0 is due to Zolotarev (see Theorem 2.5.2 [103]): for all integer $N\geq 1$,

$$\big(2\pi(1-\tfrac{1}{\gamma})\big)^{\frac{1}{2}}\, x^{\frac{\gamma+1}{2}}\, e^{1/x^{\gamma-1}} s_\gamma\big((\gamma-1)x\big) = 1+\sum_{1\leq n<N} S_n\, x^{n(\gamma-1)} + \mathcal{O}_{N,\gamma}\big(x^{N(\gamma-1)}\big), \quad \text{as } x\to 0. \tag{1.98}$$

Here $\mathcal{O}_{N,\gamma}$ means that the expansion depends on N and γ. Next note that S_n depends on n and γ but we skip the dependence in γ to simplify notation.

Remark 11. *In the Brownian case where $\gamma=2$, it is well-known that*

$$s_2(x) = \pi^{-\frac{1}{2}} x^{-\frac{3}{2}} e^{-1/x}, \quad x\in\mathbb{R}_+$$

Then, $S_0=1$ and $S_n=0$, for all $n\geq 1$. □

For generic $\gamma\in(1,2)$, this asymptotic expansion does not yield a converging power expansion (although it is the case if $\gamma=2$). See Chapter 3, Section 3.4.1 for more details on s_γ. To state our result we first need to introduce an auxiliary function derived from s_γ as follows.

Proposition 8 (Proposition 3.4)**.** *Let $\gamma\in(1,2]$. Recall from (1.97) the definition of s_γ. We introduce the following function:*

$$\forall x\in(0,\infty), \quad \theta(x) := (\gamma-1)\, x^{-1} s_\gamma(x) - \tfrac{\gamma-1}{\gamma} x^{-1-\frac{1}{\gamma}} \int_0^x dy\, y^{\frac{1}{\gamma}-1} s_\gamma(y)\ . \tag{1.99}$$

Then, the following holds true.

(i) θ is well-defined, continuous,

$$\int_0^\infty dx\, |\theta(x)| < \infty \quad \textit{and} \quad \int_0^\infty dx\, e^{-\lambda x}\theta(x) = \lambda^{\frac{1}{\gamma}} e^{-\gamma\lambda^{\frac{\gamma-1}{\gamma}}}, \quad \lambda\in(0,\infty). \tag{1.100}$$

(ii) Recall from (1.98) the definition of the sequence $(S_n)_{n\geq 0}$, with $S_0=1$. Let $(V_n)_{n\geq 0}$ be a sequence of real numbers recursively defined by $V_0=1$ and

$$\forall n\in\mathbb{N}, \qquad V_{n+1} = S_{n+1} + \big(n-\tfrac{1}{2}-\tfrac{1}{\gamma-1}\big) S_n - \big(n-\tfrac{1}{2}-\tfrac{1}{\gamma}\big) V_n\ . \tag{1.101}$$

Then, for all integer $N\geq 1$,

$$\big(2\pi(1-\tfrac{1}{\gamma})\big)^{\frac{1}{2}}\, x^{\frac{\gamma+3}{2}}\, e^{1/x^{\gamma-1}} \theta\big((\gamma-1)x\big) = 1+\sum_{1\leq n<N} V_n\, x^{n(\gamma-1)} + \mathcal{O}_{N,\gamma}\big(x^{N(\gamma-1)}\big), \tag{1.102}$$

as $x\to 0$.

We use θ to get the asymptotic expansion of the law of the total height of the normalized γ-stable tree as follows.

Theorem 9 (Theorem 3.5)**.** *Let $\gamma\in(1,2]$. We introduce the following function:*

$$\forall r\in(0,\infty), \quad \xi(r) := r^{-\frac{\gamma+1}{\gamma-1}} \theta\big(r^{-\frac{\gamma}{\gamma-1}}\big)\ . \tag{1.103}$$

where θ is defined in (1.99). Then, there exists a real valued sequence $(\beta_n)_{n\geq 1}$ and $x_1\in(0,1)$ such that

$$\sum_{n\geq 1} |\beta_n| x_1^n < \infty \quad \textit{and} \quad \forall r\in(0,\infty), \quad \sum_{n\geq 1} |\beta_n| \sup_{s\in[r,\infty)} |\xi(ns)| < \infty\ , \tag{1.104}$$

and such that

$$\forall r \in (0,\infty), \quad c_\gamma \mathbf{N}_{\mathrm{nr}}\big(\Gamma > r\big) = \sum_{n\geq 1} \beta_n \, \xi(nr) \;, \tag{1.105}$$

where we recall from (1.35) that $1/c_\gamma = \gamma \Gamma_e\big(\frac{\gamma-1}{\gamma}\big)$*,* Γ_e *standing for Euler's gamma function. Moreover, for all integers* $N \geq 1$*, as* $r \to \infty$*,*

$$\frac{1}{C_1} r^{-1-\frac{\gamma}{2}} e^{r^\gamma} \mathbf{N}_{\mathrm{nr}}\Big(\Gamma > r(\gamma-1)^{-\frac{\gamma-1}{\gamma}}\Big) = 1 + \sum_{1\leq n<N} V_n \, r^{-n\gamma} + \mathcal{O}_{N,\gamma}\big(r^{-N\gamma}\big) \;, \tag{1.106}$$

where $C_1 := (2\pi)^{-\frac{1}{2}}(\gamma-1)^{\frac{1}{2}+\frac{1}{\gamma}}\gamma^{\frac{3}{2}}\,\Gamma_e(\frac{\gamma-1}{\gamma})\exp(C_0)$*, where*

$$C_0 := \gamma \int_1^\infty \frac{du}{(u+1)^\gamma - 1} - \int_0^1 \frac{du}{u} \, \frac{(u+1)^\gamma - 1 - \gamma u}{(u+1)^\gamma - 1} \;, \tag{1.107}$$

and where the sequence $(V_n)_{n\geq 1}$ *is recursively defined by (1.101) in Proposition 8.*

Remark 12. *The convergence in (1.105) is rapid. Indeed, by (1.102), we see that* $\xi(nr)$ *is of order*

$$(nr)^{1+\frac{\gamma}{2}} \exp(-n^\gamma(\gamma-1)^{\frac{1}{\gamma-1}} r^\gamma) \;.$$

Then, the asymptotic expansion (1.106) is that of the first term of (1.105) that is $c_\gamma^{-1}\beta_1\,\xi(r)$*.* □

Remark 13. *The definition of the sequence* $(\beta_n)_{n\geq 0}$ *is involved: see Lemma 3.24 and its proof for a precise definition. However, in the Brownian case, everything can be explicitly computed: for all* $n \geq 1$*,* $\beta_n = 2$*,* $\xi(r) = (4\pi)^{-\frac{1}{2}}(2r^2-1)e^{-r^2}$*,* $c_2 = (4\pi)^{-\frac{1}{2}}$*, and we recover (1.95) from (1.105); moreover,* $C_0 = \log 2$*,* $C_1 = 4$*,* $V_0 = 1$*,* $V_1 = -\frac{1}{2}$ *and* $V_n = 0$*, for all* $n \geq 2$*.* □

To state the result concerning the diameter, we need precise results on the derivative of the $\frac{\gamma-1}{\gamma}$-stable density.

Proposition 10 (Proposition 3.6)**.** *Let* $\gamma \in (1,2]$*. Recall from (1.97) the definition of the density* s_γ*. Then* s_γ *is* C^1 *on* $\mathbb{R}_+$*,*

$$\int_0^\infty dx \, |s'_\gamma(x)| < \infty \qquad \text{and} \qquad \int_0^\infty dx \, e^{-\lambda x} s'_\gamma(x) = \lambda e^{-\gamma\lambda^{\frac{\gamma-1}{\gamma}}}, \quad \lambda \in (0,\infty). \tag{1.108}$$

Moreover, s'_γ *has the following asymptotic expansion: recall from (1.98) the definition of the sequence* $(S_n)_{n\geq 0}$*, with* $S_0 = 1$*; let* $(T_n)_{n\geq 0}$ *be a sequence of real numbers recursively defined by* $T_0 = 1$ *and*

$$\forall n \in \mathbb{N}, \qquad T_{n+1} := S_{n+1} + \big(n - \tfrac{1}{2} - \tfrac{1}{\gamma-1}\big) S_n \;. \tag{1.109}$$

Then, for all positive integers N*, we have*

$$\big(2\pi(1-\tfrac{1}{\gamma})\big)^{\frac{1}{2}} x^{\frac{3\gamma+1}{2}} e^{1/x^{\gamma-1}} s'_\gamma\big((\gamma-1)x\big) = 1 + \sum_{1\leq n<N} T_n \, x^{n(\gamma-1)} + \mathcal{O}_{N,\gamma}\big(x^{N(\gamma-1)}\big), \tag{1.110}$$

as $x \to 0$*.*

The asymptotic expansion of the law of the diameter of the normalized γ-stable tree is then given in the following theorem.

Theorem 11 (Theorem 3.7)**.** *Let $\gamma \in (1,2]$. Recall from (1.103) the definition of the function ξ. We also introduce the following function:*

$$\forall r \in (0,\infty), \quad \overline{\xi}(r) := r^{-\frac{\gamma+1}{\gamma-1}} s'_\gamma\big(r^{-\frac{\gamma}{\gamma-1}}\big) , \tag{1.111}$$

where s'_γ is the derivative of the density s_γ defined in (1.97). Then there exist two real valued sequences $(\gamma_n)_{n\geq 2}$ and $(\delta_n)_{n\geq 2}$ and $x_2 \in (0,1)$ such that

$$\sum_{n\geq 2} (|\gamma_n| + |\delta_n|) x_2^n < \infty \quad \text{and} \quad \forall r \in (0,\infty), \quad \sum_{n\geq 2} |\gamma_n| \sup_{s\in[r,\infty)} |\overline{\xi}(ns)| + |\delta_n| \sup_{s\in[r,\infty)} |\xi(ns)| < \infty , \tag{1.112}$$

and such that

$$\forall r \in (0,\infty), \quad c_\gamma \mathbf{N}_{\mathrm{nr}}\big(D > 2r\big) = \sum_{n\geq 2} \gamma_n \overline{\xi}(nr) + \delta_n \xi(nr) , \tag{1.113}$$

where we recall from (1.35) that $1/c_\gamma = \gamma \Gamma_e\big(\frac{\gamma-1}{\gamma}\big)$, Γ_e standing for Euler's gamma function. Moreover, for all integers $N \geq 1$, as $r \to \infty$,

$$\frac{1}{C_2} r^{-1-\frac{3\gamma}{2}} e^{r^\gamma} \mathbf{N}_{\mathrm{nr}}\Big(D > r(\gamma-1)^{-\frac{\gamma-1}{\gamma}}\Big) = 1 + \sum_{1\leq n<N} U_n r^{-n\gamma} + \mathcal{O}_{\gamma,N}\big(r^{-N\gamma}\big) , \tag{1.114}$$

where $C_2 := (8\pi)^{-\frac{1}{2}} (\gamma-1)^{\frac{3}{2}+\frac{1}{\gamma}} \gamma^{\frac{5}{2}} \Gamma_e\big(\frac{\gamma-1}{\gamma}\big) \exp(2C_0)$, where C_0 is defined by (1.107) and where the sequence $(U_n)_{n\geq 1}$ is recursively defined by $U_0 = 1$ and

$$\forall n \geq 1 , \qquad U_n = T_n - \tfrac{\gamma+1}{\gamma(\gamma-1)} V_{n-1} . \tag{1.115}$$

Here $(T_n)_{n\geq 0}$ is defined by (1.109) and $(V_n)_{n\geq 0}$ is defined by (1.101).

Remark 14. *The convergence in (1.113) is rapid. Indeed, by (1.110) and (1.102) we see that $\overline{\xi}(nr/2)$ and $\xi(nr/2)$ are of respective order*

$$(nr)^{1+\frac{3\gamma}{2}} \exp(-n^\gamma 2^{-\gamma} (\gamma-1)^{\frac{1}{\gamma-1}} r^\gamma) \quad \text{and} \quad (nr)^{1+\frac{\gamma}{2}} \exp(-n^\gamma 2^{-\gamma} (\gamma-1)^{\frac{1}{\gamma-1}} r^\gamma) .$$

Then the asymptotic expansion (1.114) is that of $c_\gamma^{-1} \gamma_2 \, \overline{\xi}(r) + c_\gamma^{-1} \delta_2 \, \xi(r)$. □

Remark 15. *The definitions of the sequences $(\gamma_n)_{n\geq 0}$ and $(\delta_n)_{n\geq 0}$ are involved: see the proof of Lemma 3.25 for a precise definition. However, in the Brownian case, everything can be computed explicitly:*

$$\forall n \geq 2, \quad \gamma_n = \tfrac{4}{3}(n^2-1), \quad \delta_n = -2(n^2-1) \quad \text{and} \quad \overline{\xi}(r) = \pi^{-\frac{1}{2}} r^2 \big(r^2 - \tfrac{3}{2}\big) e^{-r^2} ,$$

which allows to recover (1.96) from (1.113). Moreover, $C_2 = 8$, $U_0 = 1$, $U_1 = -3$, $U_2 = -\frac{3}{4}$ and $U_n = 0$, for all $n \geq 3$. □

1.3.3 Cutting and re-arranging trees

The study of random cutting on trees dates back to Meir and Moon [88] and has several variations. Here, we consider the following version which consists in cutting down a tree by iteratively removing random vertices. Given a rooted (graph) tree T, choose a uniform vertex and remove it. This splits T into several connected components. Retain the one containing the root and discard the other ones. Then keep repeating the same procedure on the remaining part until the tree is empty. Each vertex that has been picked and removed is referred to as a *cut*. Denote by $L(T)$ the total number of cuts. If T_n denotes the Cayley tree on n vertices, then Panholzer [93] has shown that

$$\frac{1}{\sqrt{n}} L(T_n) \xrightarrow{d} \mathrm{R}, \quad \text{as } n \to \infty, \tag{1.116}$$

where R is a random variable with Rayleigh distribution (of density function $xe^{-x^2/2}$ on $[0, \infty)$). Janson [72] has extended this result to the case of conditioned Galton–Watson trees with a finite-variance offspring distribution.

The convergence in (1.116) and its extension by Janson motivated a number of recent works, which address one or both of the following topics. The first one is proposed by Janson [72]. As we have seen, if ξ is a critical offspring distribution with finite variance σ^2, and τ_n denotes the Galton–Watson tree with offspring distribution ξ conditioned to have total progeny n, then

$$\frac{\sigma}{\sqrt{n}} \tau_n \underset{d,\mathrm{GH}}{\overset{n\to\infty}{\longrightarrow}} \mathcal{T}^{br}, \tag{1.117}$$

where $\mathcal{T}^{br}$ denotes the Brownian CRT, and $\to_{d,\mathrm{GH}}$ denotes the convergence in distribution with respect to the Gromov–Hausdorff topology. Comparing this with Janson's result, one might wonder whether it is possible to define a continuous cutting procedure on the limit tree $\mathcal{T}^{br}$ and a random variable which is an analog for the number of cuts such that (1.116) would follow from the convergence of the cutting procedures. This is studied in Addario-Berry, Broutin, and Holmgren [7], Abraham and Delmas [4], as well as Bertoin and Miermont [30]. It turns out that the continuous cutting procedure is closely related to the fragmentation process considered in Aldous and Pitman [11]. And the continuous analog for the number of cuts, which we denote by $L(\mathcal{T}^{br})$, is a measurable function of the fragmentation process (see (1.121) below). Then the authors in [7] and [30] show independently that

$$\frac{1}{\sigma\sqrt{n}} L(\tau_n) \xrightarrow{d} L(\mathcal{T}^{br}) \tag{1.118}$$

jointly with the convergence in (1.117). The special case for the Cayley trees is also shown in [4]. Combining (1.118) with (1.116) (recall that the Cayley tree corresponds to a conditioned Galton–Watson tree with Poisson offspring distribution), we see that $L(\mathcal{T}^{br})$ has Rayleigh distribution. In the case where the offspring distribution is in the domain of attraction of some α-stable distribution for $\alpha \in (1, 2)$, the conditioned Galton–Watson tree converges weakly to the α-stable tree (see Duquesne [49]). Then a result similar to (1.118) has been proved by Dieuleveut [46], where the continuous cutting procedure is induced by the self-similar fragmentation on the branch points defined in Miermont [90].

Another natural question arising from (1.116) is about the limit distribution. Note that the distribution of $L(\mathcal{T}^{br})$, which is Rayleigh, is also the distribution of the distance in $\mathcal{T}^{br}$ between two uniform nodes. This coincidence of distributions is explained in Addario-Berry et al. [7], using a bijection for the cuttings of the Cayley tree. Indeed, if we take the discarded subtrees from the cutting procedure of the Cayley tree T_n and connect their roots to make a path (see Figure 1.4), then the tree obtained is distributed as T_n. Moreover, the extremities of this path are two independent uniform nodes. Therefore, $L(T_n)$ has the same distribution as the number of nodes on a uniform path in T_n and the distribution of $L(\mathcal{T}^{br})$ follows easily from a weak convergence argument. Another explanation for the distribution of $L(\mathcal{T}^{br})$ is given in Bertoin and Miermont [30], where the argument is based on the duality of two self-similar fragmentations on $\mathcal{T}^{br}$. This kind of identity in distribution, saying that $L(\mathcal{T}^{br})$ is distributed as the distance between two uniform nodes can be extended to a general Lévy tree under the excursion measure. This is done by Abraham and Delmas [5].

It turns out that this kind of identity in distribution is also true for a general ICRT with a cutting procedure that we define below. However, the argument in [30] cannot apply, since an ICRT is not self-similar in general; nor can the argument in [5], which is based on the nice analytic properties of the underlying Lévy process. On the other hand, the bijection on the Cayley tree in [7] can be extended to the birthday trees, as a consequence of the Aldous–Broder Algorithm (Algorithm 2). The results on the cutting of ICRTs then follow from weak convergence arguments.

In the sequel, we first introduce the cutting procedures on the discrete and continuous models. Then we announce the main results, whose proofs are found in Chapter 4.

A cutting procedure on the discrete trees

Let $\boldsymbol{p}$ be a probability measure on $[n]$, for a natural number $n \geq 1$. Let T be a $\boldsymbol{p}$-tree as defined in (1.40). We introduce a cutting procedure on T which generalizes the previous one on the Cayley tree. It is more convenient for us to retain the portion containing a random vertex rather than the root. For this, we sample an independent vertex V of distribution $\boldsymbol{p}$. Recall that the tree T^V obtained by re-rooting at V is still a $\boldsymbol{p}$-tree. Therefore, this modification does not affect the distribution of $L(T)$. We perform the cutting procedure on T by picking each time a vertex according to the restriction of $\boldsymbol{p}$ to the remaining part. Denote by $L(T)$ the total number of cuts, and let $X_i, 1 \leq i \leq L(T)$, be the sequence of the cuts.

During this cutting procedure, we reassemble the discarded parts, which are the subtrees above X_i just before the cutting, by adding an edge between X_i and X_{i+1}, $1 \leq i \leq L(T) - 1$. The resulting tree (Figure 1.4), denoted by $\mathrm{cut}(T, V)$, has the same vertex set as the initial tree and contains a path which is composed of $X_1, X_2, \cdots, X_{L(T)} = V$. It follows that $L(T)$ is the number of vertices on this path.

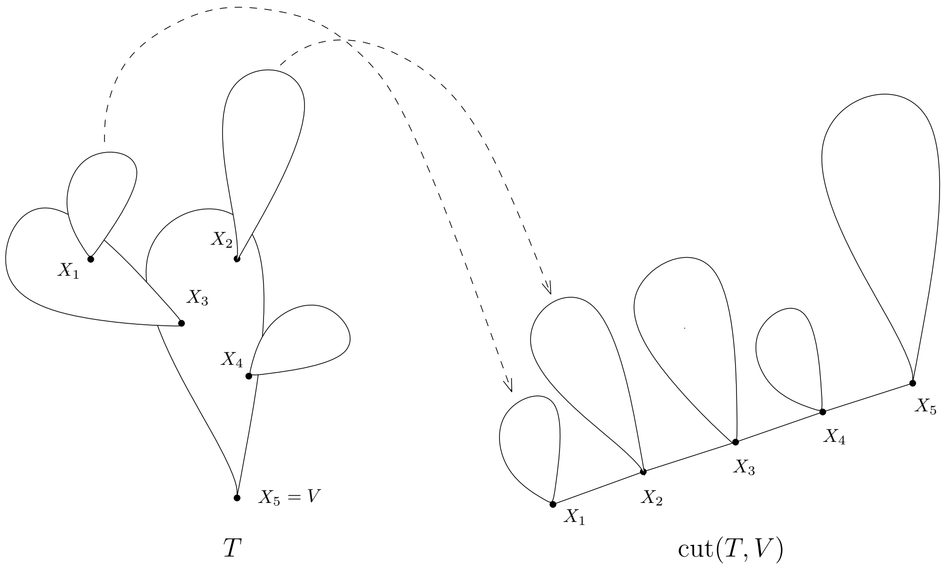

Figure 1.4 – *On the left, the cutting of T. On the right, the tree $\mathrm{cut}(T, V)$ obtained from the discarded parts of T.*

Proposition 12 (Lemma 4.10 and Proposition 4.11)**.** *For a $\boldsymbol{p}$-tree T, we have*

$$(\mathrm{cut}(T, V), V) \stackrel{d}{=} (T, V). \tag{1.119}$$

In particular, this entails

$$L(T) \stackrel{d}{=} \mathrm{Card}\{\textit{vertices on the path of } T \textit{ from the root to } V\}.$$

A cutting procedure on the continuous trees

Let $\mathcal{T}$ be an ICRT for some parameter $\boldsymbol{\theta} = (\theta_i)_{i \geq 0} \in \boldsymbol{\Theta}$, as introduced previously. Let V be a random point sampled according to the mass measure μ of $\mathcal{T}$. We mean to define a cutting procedure on $\mathcal{T}$ which is the weak limit of the previous one on the $\boldsymbol{p}$-tree T. For this, we notice that the sequence of the cuts on T can also be obtained from a Poisson point process of intensity $dt \otimes \boldsymbol{p}$ on $[0, \infty) \times T$. It suffices to filter the points of this Poisson point process in such a way that only those which fall on the part containing V are counted. Recall that $\sigma^2 = \sum_{i \geq 1} p_i^2$. Then we take $\sigma^{-1}\boldsymbol{p}$ as the discrete measure with respect to which we cut the rescaled tree σT. Under the hypothesis (1.41), we show (Proposition 4.23) that the weak limit of $(\sigma_n^{-1}\boldsymbol{p}_n, n \geq 1)$ is the following measure on $\mathcal{T}$:

$$\mathcal{L}(dx) = \theta_0^2 \ell(dx) + \sum_{i \geq 1 : \theta_i > 0} \theta_i \delta_{\beta_i}(dx), \tag{1.120}$$

where ℓ is the length measure supported on the skeleton of $\mathcal{T}$, and β_i is the ∞-degree branch point which corresponds to $\xi_{i,1}$ in the Poisson process P_i of rate $\theta_i > 0$ used in the line-breaking construction. It turns out that $\mathcal{L}$ is a σ-finite measure concentrated on $\mathrm{Sk}(\mathcal{T})$ (Lemma 4.22).

Conditional on $\mathcal{T}$, let $\mathcal{P}$ be a Poisson point process on $[0, \infty) \times \mathcal{T}$ of intensity measure $dt \otimes \mathcal{L}(dx)$. For each $t \geq 0$, define $\mathcal{P}_t = \{x \in \mathcal{T} : \exists s \leq t \text{ such that } (s, x) \in \mathcal{P}\}$. Let $\mathcal{T}_t = \mathcal{T}_t$ be the connected component of $\mathcal{T} \setminus \mathcal{P}_t$ containing V. This is the portion of the tree remaining at time t. We set $\mathcal{C} := \{t > 0 : \mu(\mathcal{T}_{t-}) > \mu(\mathcal{T}_t)\}$. Then for each $t \in \mathcal{C}$, there exists a (unique) $x(t) \in \mathcal{T}$ such that $(t, x(t)) \in \mathcal{P}$. Moreover, $x(t)$ is contained in $\mathcal{T}_{t-}$, the portion left just before the cutting.

We set

$$L(\mathcal{T}) := \int_0^\infty \mu(\mathcal{T}_s) ds, \tag{1.121}$$

which is almost surely finite (Theorem 4.4). This turns out to be the continuous analog for the number of cuts. Note that such a definition has already appeared in [7, 30] for the Brownian CRT, and in [5] for the Lévy trees.

In a similar way to the discrete case, we construct another (real) tree which partially encodes the cutting procedure on $\mathcal{T}$. For $t \in [0, \infty]$, we define $L_t := \int_0^t \mu(\mathcal{T}_s) ds$. In particular, $L_\infty = L(\mathcal{T})$. Consider the interval $[0, L_\infty]$, which is almost surely finite. For each $t \in \mathcal{C}$, graft $\mathcal{T}_{t-} \setminus \mathcal{T}_t$, the portion of the tree discarded at time t, at the point $L_t \in [0, L_\infty]$. This produces a real tree, seen as rooted at the extremity 0 of $[0, L_\infty]$. Denote by $\mathrm{cut}(\mathcal{T}, V)$ its completion. Moreover, the mass measure μ of $\mathcal{T}$ can be pushed to $\mathrm{cut}(\mathcal{T}, V)$, which yields a (possibly defective probability) measure $\hat{\mu}$ on $\mathrm{cut}(\mathcal{T}, V)$.

Let $(\boldsymbol{p}_n, n \geq 1)$ be the sequence of probability measures satisfying (1.41), and write $\mathrm{cut}(T^n, V^n)$ for the tree associated to the cutting procedure for the $\boldsymbol{p}_n$-tree T^n and the node V^n of distribution $\boldsymbol{p}_n$.

Theorem 13 (Theorem 4.4)**.** *Under* (1.41)*, we have*

$$(\sigma_n \,\mathrm{cut}(T^n, V^n), \boldsymbol{p}_n, V^n) \xrightarrow[d,\mathrm{GP}]{n \to \infty} (\mathrm{cut}(\mathcal{T}, V), \hat{\mu}, L_\infty),$$

jointly with the convergence in (1.42)*.*

Comparing this with (1.119), we obtain immediately that

Corollary 14 (Theorem 4.5, Corollary 4.6)**.** *We have the identity in distribution:*

$$(\mathrm{cut}(\mathcal{T}, V), \hat{\mu}) \stackrel{d}{=} (\mathcal{T}, \mu). \tag{1.122}$$

Moreover, $L(\mathcal{T})$ has the same distribution as the distance in $\mathcal{T}$ from the root to V, that is,

$$\mathbf{P}\big(L(\mathcal{T}) > r\big) = e^{-\frac{1}{2}\theta_0^2 r^2} \prod_{i \geq 1} (1 + \theta_i r) e^{-\theta_i r}, \quad r > 0. \tag{1.123}$$

Genealogical trees of the discrete and continuous fragmentations

In the cutting procedure described above, we only keep track of the cuts affecting the size of the connected component containing V. Following Bertoin and Miermont [30], we also consider a more general cutting procedure which keeps splitting all the connected components. It turns out that this cutting process is connected to the Aldous–Pitman fragmentation [11] for the Brownian CRT, and to a new natural fragmentation for a general ICRT.

Let $\mathcal{P}$ be the Poisson point process of intensity measure $dt \otimes \mathcal{L}(dx)$ as defined previously. For each $t \geq 0$, we obtain a "forest" from $\mathcal{T}$ by removing all the points of $\mathcal{P}_t$. More precisely, $\mathcal{P}_t$ induces an equivalence relation $\sim_t$ on $\mathcal{T}$: for $x, y \in \mathcal{T}$ we write $x \sim_t y$ if $[\![x, y]\!] \cap \mathcal{P}_t = \varnothing$, where $[\![x, y]\!]$ denotes the unique geodesic path between x and y in $\mathcal{T}$. We denote by $\mathcal{T}_x(t)$ the equivalence class containing x.

Let $\boldsymbol{\mu}^{\downarrow}(t)$ be the sequence of nonzero values of $\{\mu(\mathcal{T}_x(t)), x \in \mathcal{T}\}$ re-arranged in decreasing order. In the Brownian case ($\boldsymbol{\theta} = (1, 0, 0, \cdots)$), the process $(\boldsymbol{\mu}^{\downarrow}(t), t \geq 0)$ is exactly the Aldous–Pitman fragmentation. In the other cases, however, the process is not even Markovian because of the presence of those branch points β_i associated with the positive θ_i.

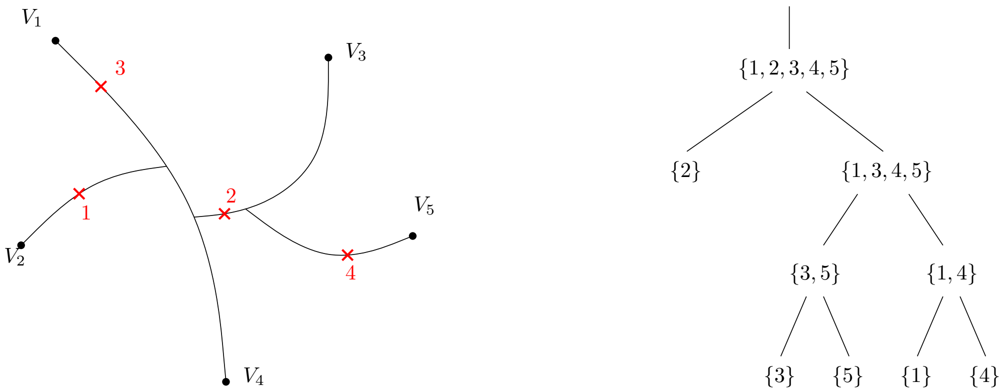

Figure 1.5 – *On the left, the subtree of $\mathcal{T}$ spanning the leaves $V_1, V_2, \dots, V_5$. The cuts which refine the partitions are represented by the crosses, and the index next to them corresponds to the order in which they appear. On the right, the genealogical tree S_5 of the partitions.*

As previously, we construct another tree to encode the cutting procedure, which can be interpreted as the genealogical tree of the fragmentation associated to the cutting. In the Brownian case, Bertoin and Miermont [30] have shown that this genealogical tree is distributed as a Brownian CRT. We extend this result to the ICRTs using a completely different method.

First, let us introduce the genealogical tree for the ICRT. Recall that from Kingman's theory [75] there is a correspondence in distribution between mass partitions and exchangeable partitions on $\mathbb{N}$, and the distribution of the latter is characterized by its restrictions on $[k], k \geq 1$. Now sample a sequence of independent points $(V_i)_{i\geq 1}$ according to the mass measure μ. Then $\sim_t$ induces an (exchangeable) partition on $\mathbb{N}$ by setting $i \sim_t j$ if $V_i \sim_t V_j$. In particular, the mass $\mu(\mathcal{T}_{V_i}(t))$ can be recovered as the asymptotic frequency:

$$\mu\big(\mathcal{T}_{V_i}(t)\big) = \lim_{k\to\infty} \frac{1}{k} \sum_{j=1}^{k} \mathbf{1}_{\{j \sim_t i\}}, \quad \text{almost surely,}$$

by the law of large numbers. We set for $t \in [0, \infty]$ and $i \geq 1$

$$L_t^i := \int_0^t \mu\big(\mathcal{T}_{V_i}(s)\big) ds.$$

As $V_i \stackrel{d}{=} V$ and $(\mathcal{T}_{V_i}(s))_{s\geq 0} \stackrel{d}{=} (\mathcal{T}_s)_{s\geq 0}$ for each $i \geq 1$, it follows that we have $L_\infty^i < \infty$ for all $i \geq 1$, almost surely. For each pair (i, j) such that $i \neq j$, let $\tau(i, j)$ be the first moment when $[\![V_i, V_j]\!]$ contains an element of $\mathcal{P}$ (or more precisely, of the projection of $\mathcal{P}$ onto $\mathcal{T}$). Then $\tau(i, j) = \tau(j, i)$ records the instant when V_i and V_j are separated into different equivalence classes. It follows from the properties of $\mathcal{T}$ and $\mathcal{P}$ that $\tau(i, j)$ is almost surely finite.

For each $k \geq 1$, we can construct a k-leafed tree S_k which represents the genealogical structure of how the partitions of $[k]$ induced by $\sim_t, t \geq 0$ evolve into singletons $\{1\}, \cdots, \{k\}$. See figure 1.5. Moreover, we equip S_k with a distance d_k satisfying that

$$d_k(\partial, \{i\}) = L^i_\infty, \quad d_k(\{i\}, \{j\}) = L^i_\infty + L^j_\infty - 2L^i_{\tau(i,j)}, \quad 1 \leq i < j \leq k, \tag{1.124}$$

where ∂ denotes the root of S_k.

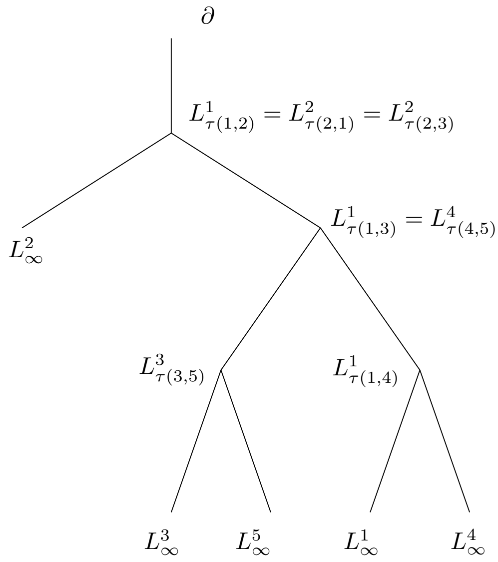

Figure 1.6 – *The same S_5 as in Figure 1.5 but equipped with a distance. The numbers here are the distances of the nodes from the root.*

We can construct $(S_k)_{k\geq 1}$ in such a way that $S_k \subset S_{k+1}$ as metric spaces. Let $\mathrm{cut}(\mathcal{T})$ be the completion of $\cup_{k\geq 1} S_k$.

Similarly, for each birthday tree T^n on $[n]$, we can define a complete cutting procedure on T^n by first generating a random permutation $(X_{n1}, X_{n2}, \cdots, X_{nn})$ on the vertex set $[n]$ and then removing X_{ni} one by one. Here the permutation $(X_{n1}, X_{n2}, \ldots, X_{nn})$ is constructed by sampling, for $i \geq 1$, X_{ni} according to the restriction of $\boldsymbol{p}_n$ to $[n] \setminus \{X_{nj}, j < i\}$. We define a new genealogy on $[n]$ by making X_{ni} an ancestor of X_{nj} if $i < j$ and X_{ni} is in the connected component containing X_{nj} when it (X_{ni}) is removed. If we denote by $\mathrm{cut}(T^n)$ the corresponding genealogical tree, then the number of the vertices in the path of $\mathrm{cut}(T^n)$ between the root X_{n1} and an arbitrary vertex v is precisely equal to the number of the cuts necessary to isolate this vertex v. See Figure 1.7 for an example of $\mathrm{cut}(T^n)$.

Theorem 15 (Theorems 4.7 and 4.8)**.** *Suppose that* (1.41) *holds. Then, we have*

$$\big(\sigma_n \,\mathrm{cut}(T^n), \boldsymbol{p}_n\big) \xrightarrow[d,\mathrm{GP}]{n\to\infty} \big(\mathrm{cut}(\mathcal{T}), \nu\big),$$

jointly with the convergence in (1.42)*. Here, ν is the weak limit of the empirical measures $\frac{1}{k}\sum_{i=0}^{k-1} \delta_i$, which exists almost surely conditional on $\mathcal{T}$. Moreover, we have*

$$(\mathrm{cut}(\mathcal{T}), \nu) \stackrel{d}{=} (\mathcal{T}, \mu). \tag{1.125}$$

Recovering the Brownian CRT from its genealogy of the cutting process

The identity in distribution (1.125) gives rise to the following question: given an ICRT $\mathcal{H}$, then $\mathcal{H}$ has the same distribution as $\mathrm{cut}(\mathcal{T})$, and one might wonder whether it is possible to construct another (real)

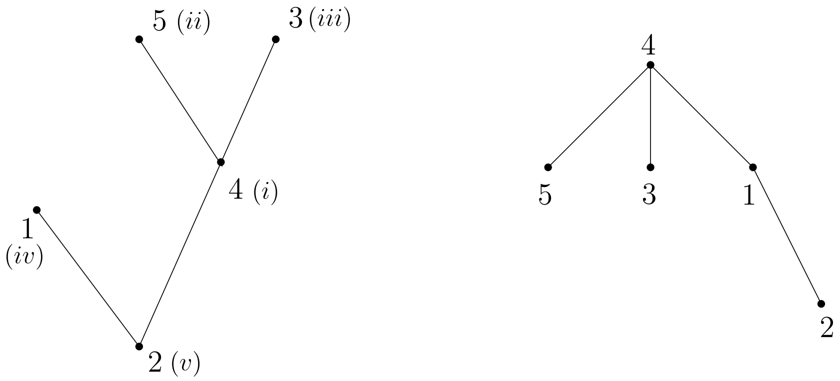

Figure 1.7 – *On the left, a cutting of T^n where the roman numbers represent the order in which the vertices are removed. On the right, the corresponding* $\mathrm{cut}(T^n)$.

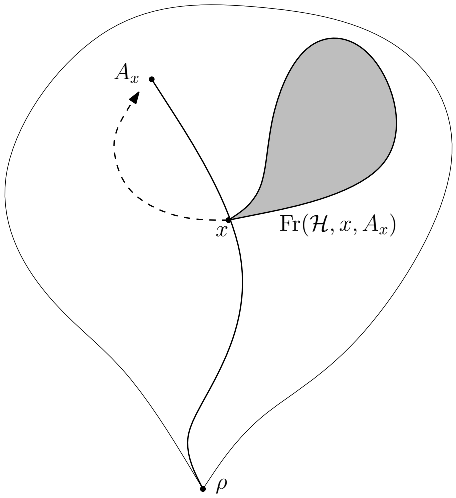

Figure 1.8 – *The surgical operation on the tree $\mathcal{H}$ for a single branch point x.*

tree $\mathcal{Q}$ such that

$$(\mathcal{Q}, \mathcal{H}) \stackrel{d}{=} (\mathcal{T}, \mathrm{cut}(\mathcal{T})). \tag{1.126}$$

In the case where $\mathcal{H}$ is a Brownian CRT, we have succeeded in constructing $\mathcal{Q} = \mathrm{shuff}(\mathcal{H})$ by the following subtree-shuffling procedure. Let ρ denote the root of $\mathcal{H}$ and $\mathrm{Br}(\mathcal{H})$ the set of branch points of $\mathcal{H}$, which is a countable set with probability 1. For each $x \in \mathcal{H}$, the subtree at x, denoted by $\mathrm{Sub}(\mathcal{H}, x)$, is the set of those points y such that $x \in [\![\rho, y]\!]$. For $x \in \mathrm{Br}(\mathcal{H})$, sample independently a random point A_x according to the restriction of $\mu_{\mathcal{H}}$, the mass measure of $\mathcal{H}$, to $\mathrm{Sub}(\mathcal{H}, x)$, and set $\mathrm{Fr}(\mathcal{H}, x, A_x)$ to be the set of those $y \in \mathrm{Sub}(\mathcal{H}, x)$ for which the closest common ancestor of y and A_x is $y \wedge A_x = x$. Detach $\mathrm{Fr}(\mathcal{H}, x, A_x)$ and re-attach it at A_x. Do this for every branch point x of $\mathcal{H}$; the points of the skeleton that are not branch points are not used (see Figure 1.8). In Chapter 5 we show that this definition makes sense and the tree obtained indeed satisfies (1.126).

Let us explain the motivation underlying this construction. First, assume that $\mathcal{H} = \mathrm{cut}(\mathcal{T}, V)$ for some Brownian CRT $\mathcal{T}$ along with the point V of distribution μ. Then it is intuitively clear how to “reverse" the construction of $\mathrm{cut}(\mathcal{T}, V)$: for each $x \in \mathrm{Br}(\mathcal{H}) \cap [0, L_\infty]$, $\mathrm{Fr}(\mathcal{H}, x, L_\infty)$ is a subtree on the interval $[0, L_\infty]$, and is the completion of $\mathcal{T}_{t-} \setminus \mathcal{T}_t$ for some $t \in \mathcal{C}$ by our construction of $\mathrm{cut}(\mathcal{T}, V)$. We detach this subtree and then re-attach it at the point A'_x, which records the location where the cut at time t falls on $\mathcal{T}$. See Figure 1.9. We show that given $\mathcal{H}$, A'_x has distribution $\mu_{\mathcal{H}}$ restricted to $\mathrm{Sub}(\mathcal{H}, x) \setminus \mathrm{Fr}(\mathcal{H}, x, L_\infty)$. Furthermore, when the attach points A'_x are sampled beforehand, the order

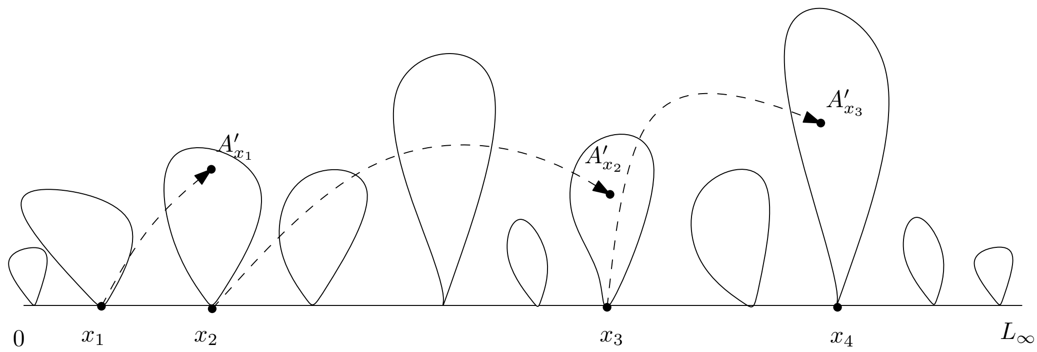

Figure 1.9 – *The detaching-reattaching operations on three branch points x_1, x_2, x_3. The order of the operations is unimportant.*

of the subtrees to be detached and reattached does not matter.

More generally, for each $k \geq 1$, we can define a k-cutting procedure on $\mathcal{T}$ which uses the elements of $\mathcal{P}$ as cuts to isolate k independent leaves $V_1, \cdots, V_k$. Similarly, this cutting procedure is (partially) encoded by $\mathrm{cut}(\mathcal{T}, V_1, \cdots, V_k)$, which is a real tree obtained by grafting discarded parts on a backbone S_k, which is no longer a path but a tree with k leaves. In an analogous way to $\mathrm{cut}(\mathcal{T}, V)$, by detaching the subtrees grafted on S_k and then re-attaching them at random points we can "reverse" the construction of $\mathrm{cut}(\mathcal{T}, V_1, \cdots, V_k)$. It should not come as a surprise that $\mathrm{cut}(\mathcal{T}, V_1, \cdots, V_k) \to \mathrm{cut}(\mathcal{T})$ as $k \to \infty$, and we prove that the sequence of "reverses" converges almost surely to a tree having all the properties that we want for the reverse of $\mathrm{cut}(\mathcal{T})$.

In fact, the "reverse" construction of $\mathrm{cut}(\mathcal{T}, V_1, \cdots, V_k)$ works not only for the Brownian CRT but for any ICRT (see Theorem 4.31 and Proposition 5.7). But the argument showing the limit exists when k tends to infinity (Theorem 5.8) relies heavily on the scaling property of the Brownian CRT. We conjecture that the "complete shuffling" described above is also the correct transformation for ICRTs.

Chapter 2

Height and Diameter of Brownian trees

The results of this chapter are from the article [100], submitted for publication.

Contents

By computations on generating functions, Szekeres proved in 1983 that the law of the diameter of a uniformly distributed rooted labelled tree with n vertices, rescaled by a factor $n^{-\frac{1}{2}}$, converges to a distribution whose density is explicit. Aldous observed in 1991 that this limiting distribution is the law of the diameter of the Brownian tree. In our article, we provide a computation of this law which is directly based on the normalized Brownian excursion. Moreover, we provide an explicit formula for the joint law of the height and diameter of the Brownian tree, which is a new result.

2.1 Introduction

For any integer $n \geq 1$, let T_n be a uniformly distributed random rooted labelled tree with n vertices and we denote by D_n its diameter with respect to the graph distance. By computations on generating functions, Szekeres [98] proved that

$$n^{-\frac{1}{2}} D_n \xrightarrow{\text{(law)}} \Delta\ , \tag{2.1}$$

where Δ is a random variable whose probability density f_Δ is given by

$$f_\Delta(y) = \frac{\sqrt{2\pi}}{3} \sum_{n\geq 1} \left(\frac{64}{y^4}(4b_{n,y}^4 - 36b_{n,y}^3 + 75b_{n,y}^2 - 30b_{n,y}) + \frac{16}{y^2}(2b_{n,y}^3 - 5b_{n,y}^2) \right) e^{-b_{n,y}}, \tag{2.2}$$

where $b_{n,y} := 8(\pi n/y)^2$, for all $y \in (0,\infty)$ and for all integers $n \geq 1$. This result is implicitly written in Szekeres [98] p. 395 formula (12). See also Broutin and Flajolet [38] for a similar result for binary trees. On the other hand, Aldous [8, 10] has proved that T_n, whose graph distance is rescaled by a factor $n^{-\frac{1}{2}}$, converges in distribution to the Brownian tree (also called Continuum Random Tree) that is a random compact metric space. From this, Aldous has deduced that Δ has the same distribution as the diameter of the Brownian tree: see [9], Section 3.4, (though formula (41) there is not accurate). As proved by Aldous [10] and by Le Gall [80], the Brownian tree is coded by the normalized Brownian excursion of length 1 (see below for more details). Then, the question was raised by Aldous [9] that whether we can establish (2.2) directly from computations on the normalized Brownian excursion. In this work, we present a solution to this question: we compute the Laplace transform for the law of the diameter of the Brownian tree based on Williams' decomposition of Brownian excursions. We also provide a formula for the joint law of the total height and diameter of the Brownian tree, which appears to be new. Before stating precisely our results, let us first recall the definition of the Brownian tree coded by the normalized Brownian excursion.

Normalized Brownian excursion. Let $X=(X_t)_{t\geq 0}$ be a continuous process defined on a probability space $(\Omega,\mathcal{F},\mathbf{P})$ such that $(\frac{1}{\sqrt{2}}X_t)_{t\geq 0}$ is distributed as a linear standard Brownian motion such that $\mathbf{P}(X_0=0)=1$ (the reason for the normalizing constant $\sqrt{2}$ is explained below). Thus,

$$\forall u\in\mathbb{R},\ t\in\mathbb{R}_+,\quad \mathbf{E}\big[e^{iuX_t}\big]=e^{-tu^2}\,.$$

For all $t\in[0,\infty)$, we set $I_t=\inf_{s\in[0,t]}X_s$. Then, the reflected process $X-I$ is a strong Markov process, the state 0 is instantaneous in $(0,\infty)$ and recurrent, and $-I$ is a local time at level 0 for $X-I$ (see Bertoin [22], Chapter VI). We denote by $\mathbf{N}$ the *excursion measure* associated with the local time $-I$; $\mathbf{N}$ is a sigma finite measure on the space of continuous paths $\mathbb{C}(\mathbb{R}_+,\mathbb{R}_+)$. More precisely, let $\bigcup_{i\in\mathcal{I}}(a_i,b_i)=\big\{t>0: X_t-I_t>0\big\}$ be the excursion intervals of the reflected process $X-I$ above 0; for all $i\in\mathcal{I}$, we set $e_i(s)=X_{(a_i+s)\wedge b_i}-I_{a_i}$, $s\in\mathbb{R}_+$. Then,

$$\sum_{i\in\mathcal{I}}\delta_{(-I_{a_i},e_i)}\ \text{is a Poisson point measure on } \mathbb{R}_+\times\mathbb{C}(\mathbb{R}_+,\mathbb{R}_+)\text{ with intensity } dt\,\mathbf{N}(de). \tag{2.3}$$

We shall denote by $e=(e_t)_{t\geq 0}$ the canonical process on $\mathbb{C}(\mathbb{R}_+,\mathbb{R}_+)$. We define its *lifetime* by

$$\zeta=\sup\{t\geq 0: e_t>0\}\,, \tag{2.4}$$

with the convention that $\sup\varnothing=0$. Then, $\mathbf{N}$-a.e. $e_0=0$, $\zeta\in(0,\infty)$ and for all $t\in(0,\zeta)$, $e_t>0$. Moreover, one has

$$\forall\lambda\in\mathbb{R}_+,\qquad \mathbf{N}\big(1-e^{-\lambda\zeta}\big)=\sqrt{\lambda}\quad\text{and}\quad \mathbf{N}\big(\zeta\in dr\big)=\frac{dr}{2\sqrt{\pi}\,r^{3/2}}\,. \tag{2.5}$$

See Blumenthal [34] IV.1 for more detail.

Let us briefly recall the scaling property of e under $\mathbf{N}$. To that end, recall that X satisfies the following scaling property: for all $r\in(0,\infty)$, $(r^{-\frac{1}{2}}X_{rt})_{t\geq 0}$ has the same law as X, which easily entails that

$$\big(r^{-\frac{1}{2}}e_{rt}\big)_{t\geq 0}\quad\text{under}\quad r^{\frac{1}{2}}\,\mathbf{N}\quad\overset{\text{(law)}}{=}\quad e\quad\text{under}\quad\mathbf{N}\,. \tag{2.6}$$

This scaling property implies that there exists a family of laws on $\mathbb{C}(\mathbb{R}_+,\mathbb{R}_+)$ denoted by $\mathbf{N}(\,\cdot\,|\,\zeta=r)$, $r\in(0,\infty)$, such that $r\mapsto\mathbf{N}(\,\cdot\,|\,\zeta=r)$ is weakly continous on $\mathbb{C}(\mathbb{R}_+,\mathbb{R}_+)$, such that $\mathbf{N}(\,\cdot\,|\,\zeta=r)$-a.s. $\zeta=r$ and such that

$$\mathbf{N}=\int_0^\infty \mathbf{N}(\,\cdot\,|\,\zeta=r)\,\mathbf{N}\big(\zeta\in dr\big)\,. \tag{2.7}$$

Moreover, by (2.6), $\big(r^{-\frac{1}{2}}e_{rt}\big)_{t\geq 0}$ under $\mathbf{N}(\,\cdot\,|\,\zeta=r)$ has the same law as e under $\mathbf{N}(\,\cdot\,|\,\zeta=1)$. To simplify notation we set

$$\mathbf{N}_{\text{nr}}:=\mathbf{N}(\,\cdot\,|\,\zeta=1)\,. \tag{2.8}$$

Thus, for all measurable functions $F:\mathbb{C}(\mathbb{R}_+,\mathbb{R}_+)\to\mathbb{R}_+$,

$$\mathbf{N}\big[F(e)\big]=\frac{1}{2\sqrt{\pi}}\int_0^\infty dr\,r^{-\frac{3}{2}}\,\mathbf{N}_{\text{nr}}\Big[F\Big(\big(r^{\frac{1}{2}}e_{t/r}\big)_{t\geq 0}\Big)\Big]\,. \tag{2.9}$$

Remark 16. *The standard Ito measure* $\mathbf{N}_{\mathrm{Ito}}^{+}$ *of positive excursions, as defined for instance in Revuz & Yor [97] Chapter XII Theorem 4.2, is derived from* $\mathbf{N}$ *by the following scaling relations:*

$$\mathbf{N}_{\mathrm{Ito}}^{+} \textit{ is the law of } \tfrac{1}{\sqrt{2}} e \textit{ under } \tfrac{1}{\sqrt{2}}\mathbf{N} \textit{ and thus, } \mathbf{N}_{\mathrm{Ito}}^{+}(\,\cdot\,|\,\zeta=1) \textit{ is the law of } \tfrac{1}{\sqrt{2}} e \textit{ under } \mathbf{N}_{\mathrm{nr}}.$$

Consequently, the law $\mathbf{N}_{\mathrm{nr}}$ *is not the standard version for normalized Brownian excursion measure. However, we shall refer to it as the normalized Brownian excursion measure.* □

Real trees. Let us recall the definition of *real trees* that are metric spaces generalizing graph-trees: let (T, d) be a metric space; it is a real tree if the following statements hold true.

(a) For all $\sigma_1, \sigma_1 \in T$, there is a unique isometry $f : [0, d(\sigma_1, \sigma_2)] \to T$ such that $f(0) = \sigma_1$ and $f(d(\sigma_1, \sigma_2)) = \sigma_2$. In this case, we set $[\![\sigma_1, \sigma_2]\!] := f([0, d(\sigma_1, \sigma_2)])$.

(b) For any continuous injective function $q : [0, 1] \to T$, $q([0, 1]) = [\![q(0), q(1)]\!]$.

When a point $\rho \in T$ is distinguished, (T, d, ρ) is said to be a *rooted* real tree, ρ being the *root* of T. Among connected metric spaces, real trees are characterized by the so-called four-point inequality: we refer to Evans [57] or to Dress, Moulton & Terhalle [47] for a detailed account on this property. Let us briefly mention that the set of (pointed) isometry classes of compact rooted real trees can be equipped with the (pointed) Gromov–Hausdorff distance which makes it into a Polish space: see Evans, Pitman & Winter [59], Theorem 2, for more detail on this intrinsic point of view that we do not adopt here.

Coding of real trees. Real trees can be constructed through continuous functions. Recall that e stands for the canonical process on $\mathbb{C}(\mathbb{R}_+, \mathbb{R}_+)$. We assume here that e has a compact support, that $e_0 = 0$ and that e is not identically null; recall from (2.4) the definition of its lifetime ζ. Then, our assumptions on e entail that $\zeta \in (0, \infty)$. For $s, t \in [0, \zeta]$, we set

$$b(s,t) := \inf_{r \in [s \wedge t, s \vee t]} e_r \quad \text{and} \quad d(s,t) := e_t + e_s - 2b(s,t) \; .$$

It is easy to see that d is a pseudo-distance on $[0, \zeta]$. We define the equivalence relation $\sim$ by setting $s \sim t$ iff $d(s, t) = 0$; then we set

$$\mathcal{T} := [0, \zeta]/\sim \; . \tag{2.10}$$

The function d induces a distance on the quotient set $\mathcal{T}$ that we keep denoting d for simplicity. We denote by $p : [0, \zeta] \to \mathcal{T}$ the canonical projection. Clearly p is continuous, which implies that $(\mathcal{T}, d)$ is a compact metric space. Moreover, it is shown that $(\mathcal{T}, d)$ is a real tree (see Duquesne & Le Gall [52], Theorem 2.1, for a proof). We take $\rho = p(0)$ as the *root* of $\mathcal{T}$. The *total height* and the *diameter* of $\mathcal{T}$ are thus given by

$$\Gamma = \max_{\sigma \in \mathcal{T}} d(\rho, \sigma) = \max_{t \geq 0} e_t \quad \text{and} \quad D = \max_{\sigma, \sigma' \in \mathcal{T}} d(\sigma, \sigma') = \max_{s,t \geq 0} \big(e_t + e_s - 2b(s,t) \big). \tag{2.11}$$

We also define on $\mathcal{T}$ a finite measure m called the *mass measure* that is the pushforward measure of the Lebesgue measure on $[0, \zeta]$ by the canonical projection p. Namely, for all continuous functions $f : \mathcal{T} \to \mathbb{R}_+$,

$$\int_{\mathcal{T}} f(\sigma)\, \mathrm{m}(d\sigma) = \int_0^{\zeta} f(p(t))\, dt \; . \tag{2.12}$$

Note that $\mathrm{m}(\mathcal{T}) = \zeta$.

Brownian tree. The random rooted compact real tree $(\mathcal{T}, d, \rho)$ coded by e under the normalized Brownian excursion measure $\mathbf{N}_{\mathrm{nr}}$ defined in (2.8) is the *Brownian tree*. Here, we recall some properties of the Brownian tree. To that end, for any $\sigma \in \mathcal{T}$, we denote by $\mathrm{n}(\sigma)$ the number of connected components of the open set $\mathcal{T}\backslash\{\sigma\}$. Note that $\mathrm{n}(\sigma)$ is possibly infinite. We call this number the *degree* of σ. We say that σ is a *branch point* if $\mathrm{n}(\sigma) \geq 3$ and that σ is a *leaf* if $\mathrm{n}(\sigma) = 1$. We denote by $\mathtt{Lf}(\mathcal{T}) := \{\sigma \in \mathcal{T} : \mathrm{n}(\sigma) = 1\}$ the *set of leaves* of $\mathcal{T}$. Then the following holds true:

$$\mathbf{N}_{\mathrm{nr}}\text{-a.s.} \quad \forall \sigma \in \mathcal{T}, \quad \mathrm{n}(\sigma) \in \{1, 2, 3\}, \quad \mathrm{m} \text{ is diffuse} \quad \text{and} \quad \mathrm{m}\big(\mathcal{T}\backslash\mathtt{Lf}(\mathcal{T})\big) = 0\,, \tag{2.13}$$

where we recall from (2.12) that m stands for the mass measure. The Brownian tree has therefore only binary branch points. The fact that the mass measure is diffuse and supported by the set of leaves makes the Brownian tree a *continuum random tree* according to Aldous' terminology (see Aldous [10]). For more detail on (2.13), see for instance Duquesne & Le Gall [52].

The choice of the normalizing constant $\sqrt{2}$ for the underlying Brownian motion X is motivated by the following fact: let T_n^* be uniformly distributed on the set of rooted *planar* trees with n vertices. We view T_n^* as a graph embedded in the clockwise oriented upper half-plane, whose edges are segments of unit length and whose root is at the origin. Let us consider a particle that explores T_n^* as follows: its starts at the root and then it moves continuously on the tree at unit speed from the left to the right, backtracking as less as possible. During this exploration the particle visits each edge exactly twice and its journey lasts $2(n-1)$ units of time. For all $t \in [0, 2(n-1)]$, we denote by $C_t^{(n)}$ the distance between the root and the position of the particle at time t. The process $(C_t^{(n)})_{t\in[0,2(n-1)]}$ is called the *contour process* of T_n^*. Following an idea of Dwass [56], we can check that the contour process $(C_t^{(n)})_{t\in[0,2(n-1)]}$ is distributed as the (linear interpolation of the) simple random walk starting from 0, conditioned to stay nonnegative on $[0, 2(n-1)]$ and conditioned to hit the value -1 at time $2n-1$. Using a variant of Donsker's invariance principle, the rescaled contour function $(n^{-\frac{1}{2}} C^{(n)}_{2(n-1)t})_{t\in[0,1]}$ converges in law towards e under $\mathbf{N}_{\mathrm{nr}}$: see for instance Le Gall [82]. Thus,

$$n^{-\frac{1}{2}} D_n^* \xrightarrow{\text{(law)}} D \quad \text{under} \quad \mathbf{N}_{\mathrm{nr}},$$

where D_n^* stands for the diameter of T_n^* and D is the diameter of the Brownian tree given by (2.11).

Remark 17. *In the first paragraph of Introduction, we introduce the random tree T_n, which is uniformly distributed on the set of rooted labelled trees with n vertices. The law of T_n is therefore distinct from that of T_n^*, which is uniformly distributed on the set of rooted planar trees with n vertices. Aldous [10] has proved that the tree T_n, whose graph distance is rescaled by a factor $n^{-\frac{1}{2}}$, converges to the tree coded by $\sqrt{2}\mathrm{e}$ under $\mathbf{N}_{nr}$. Thus,*

$$\Delta \overset{\text{(law)}}{=} \sqrt{2} D \quad \textit{under} \quad \mathbf{N}_{nr}\,. \tag{2.14}$$

See Remark 19 below. □

In this article, we prove the following result that characterizes the joint law of the height and diameter of the Brownian tree.

Theorem 2.1. *Recall from (2.8) the definition of $\mathbf{N}_{\mathrm{nr}}$ and recall from (2.11) the definitions of Γ and D. We set*

$$\forall \lambda, y, z \in (0, \infty), \quad \mathrm{L}_\lambda(y, z) := \frac{1}{2\sqrt{\pi}} \int_0^\infty e^{-\lambda r} r^{-\frac{3}{2}}\, \mathbf{N}_{\mathrm{nr}}\big(r^{\frac{1}{2}} D > 2y\,;\, r^{\frac{1}{2}}\Gamma > z\big)\, dr\,. \tag{2.15}$$

Note that

$$\forall \lambda, y, z \in (0, \infty), \quad \mathrm{L}_1(y, z) = \lambda^{-\frac{1}{2}} \mathrm{L}_\lambda\big(\lambda^{-\frac{1}{2}} y\,,\, \lambda^{-\frac{1}{2}} z\big)\,. \tag{2.16}$$

Then,

$$\mathrm{L}_1(y,z) = \coth(y\vee z) - 1 - \tfrac{1}{4}\mathbf{1}_{\{z\leq 2y\}}\frac{\sinh(2q)-2q}{\sinh^4(y)}\ , \tag{2.17}$$

where $q = y\wedge(2y-z)$*. In particular, this implies that*

$$\forall \lambda, z\in(0,\infty), \quad \mathrm{L}_\lambda(0,z) = \sqrt{\lambda}\coth(z\sqrt{\lambda}) - \sqrt{\lambda} \tag{2.18}$$

and

$$\forall \lambda, y\in(0,\infty), \quad \mathrm{L}_\lambda(y,0) = \sqrt{\lambda}\coth(y\sqrt{\lambda}) - \sqrt{\lambda} - \sqrt{\lambda}\,\frac{\sinh(2y\sqrt{\lambda})-2y\sqrt{\lambda}}{4\sinh^4(y\sqrt{\lambda})}\ . \tag{2.19}$$

Corollary 2.2. *For all* $y,z\in(0,\infty)$*, we set*

$$\rho = z\vee \tfrac{y}{2} \quad \text{ and } \quad \delta = \big(\tfrac{2(y-z)}{y}\vee 0\big)\wedge 1\ . \tag{2.20}$$

Then we have

$$\mathbf{N}_{\mathrm{nr}}\big(D>y\,;\,\Gamma>z\big) = 2\sum_{n\geq 1}\big(2n^2\rho^2-1\big)\,e^{-n^2\rho^2}+ \tag{2.21}$$
$$\frac{1}{6}\sum_{n\geq 2}n(n^2-1)\Big[\big[(n+\delta)^2y^2-2\big]e^{-\frac{1}{4}(n+\delta)^2y^2}-\big[(n-\delta)^2y^2-2\big]e^{-\frac{1}{4}(n-\delta)^2y^2}+\delta y(n^3y^3-6ny)e^{-\frac{1}{4}n^2y^2}\Big]$$

and

$$\begin{aligned}\mathbf{N}_{\mathrm{nr}}\big(D\leq y\,;\,\Gamma\leq z\big) = {} & \frac{4\pi^{5/2}}{\rho^3}\sum_{n\geq 1}n^2e^{-n^2\pi^2/\rho^2} - \\ & \frac{32\pi^{3/2}}{3}\sum_{n\geq 1}n\sin(2\pi n\delta)\Big(\frac{2}{y^5}(2a_{n,y}^2-9a_{n,y}+6)-\frac{3\delta^2-1}{y^3}(a_{n,y}-1)\Big)e^{-a_{n,y}}+ \\ & \frac{16\pi^{1/2}}{3}\sum_{n\geq 1}\delta\cos(2\pi n\delta)\Big(\frac{1}{y^3}(6a_{n,y}^2-15a_{n,y}+3)-\frac{\delta^2-1}{2y}a_{n,y}\Big)e^{-a_{n,y}}+ \\ & \frac{16\pi^{1/2}}{3}\sum_{n\geq 1}\delta\Big(\frac{1}{y^3}(4a_{n,y}^3-24a_{n,y}^2+27a_{n,y}-3)+\frac{1}{2y}(2a_{n,y}^2-3a_{n,y})\Big)e^{-a_{n,y}},\end{aligned} \tag{2.22}$$

where we set $a_{n,y} = 4(\pi n/y)^2$ *for all* $y\in(0,\infty)$ *and for all* $n\geq 1$ *to simplify notation. In particular,* (2.21) *implies*

$$\mathbf{N}_{\mathrm{nr}}\big(\Gamma>y\big) = 2\sum_{n\geq 1}\big(2n^2y^2-1\big)\,e^{-n^2y^2}, \tag{2.23}$$

and

$$\mathbf{N}_{\mathrm{nr}}\big(D>y\big) = \sum_{n\geq 2}(n^2-1)\big(\tfrac{1}{6}n^4y^4-2n^2y^2+2\big)e^{-n^2y^2/4}. \tag{2.24}$$

On the other hand, (2.22) *implies*

$$\mathbf{N}_{\mathrm{nr}}\big(\Gamma\leq y\big) = \frac{4\pi^{5/2}}{y^3}\sum_{n\geq 1}n^2e^{-n^2\pi^2/y^2}, \tag{2.25}$$

and

$$\mathbf{N}_{\mathrm{nr}}\big(D\leq y\big) = \frac{\sqrt{\pi}}{3}\sum_{n\geq 1}\Big(\frac{8}{y^3}(24a_{n,y}-36a_{n,y}^2+8a_{n,y}^3)+\frac{16}{y}a_{n,y}^2\Big)e^{-a_{n,y}}\ . \tag{2.26}$$

Thus the law of D under $\mathbf{N}_{\mathrm{nr}}$ has the following density:

$$f_D(y) = \frac{1}{12}\sum_{n\geq 1}\left(n^8y^5 - n^6y^3(20+y^2) + 20n^4y(3+y^2) - 60n^2y\right)e^{-n^2y^2/4} \tag{2.27}$$

$$= \frac{2\sqrt{\pi}}{3}\sum_{n\geq 1}\Big(\frac{16}{y^4}(4a_{n,y}^4 - 36a_{n,y}^3 + 75a_{n,y}^2 - 30a_{n,y}) + \frac{8}{y^2}(2a_{n,y}^3 - 5a_{n,y}^2)\Big)e^{-a_{n,y}} . \tag{2.28}$$

Remark 18. *We derive (2.22) from (2.21) using the following identity on the theta function due to Jacobi (1828), which is a consequence of Poisson summation formula:*

$$\forall t\in(0,\infty), \forall x,y\in\mathbb{C}, \quad \sum_{n\in\mathbb{Z}} e^{-(x+n)^2t-2\pi iny} = e^{2\pi ixy}\left(\frac{\pi}{t}\right)^{\frac{1}{2}}\sum_{n\in\mathbb{Z}} e^{-\frac{\pi^2(y+n)^2}{t}+2\pi inx} . \tag{2.29}$$

See for instance Weil [101], Chapter VII, Equation (12). Not surprisingly, (2.29) can also be used to derive (2.25) from (2.23), to derive (2.26) from (2.24), or to derive (2.28) from (2.27). □

Remark 19. *We obtain (2.27) (resp. (2.28)) by differentiating (2.24) (resp. (2.26)). By (2.14), we have*

$$\forall y\in(0,\infty), \quad f_\Delta(y) = \frac{1}{\sqrt{2}}f_D\Big(\frac{y}{\sqrt{2}}\Big) ,$$

which immediately entails (2.2) from (2.28), since $a_{n,y/\sqrt{2}} = 8(\pi n/y)^2 = b_{n,y}$. □

Remark 20. *Recall that $\Gamma = \max_{t\geq 0} e_t$. Equations (2.23) and (2.25) are consistent with previous results on the distribution of the maximum of Brownian excursion: see for example Chung [45], though we need to keep in mind the difference between $\mathbf{N}_{\mathrm{nr}}$ and $\mathbf{N}^+_{\mathrm{Ito}}$, as explained in Remark 16.* □

2.2 Preliminaries

A geometric property on diameters of real trees. We begin with a simple observation on the total height and diameter of a real tree.

Lemma 2.3. *Let (T, d, ρ) be a compact rooted real tree. Then $\Gamma\leq D\leq 2\Gamma$, where*

$$\Gamma = \sup_{u\in\mathrm{T}} d(u,\rho) \quad \textit{and} \quad D = \sup_{u,v\in\mathrm{T}} d(u,v) .$$

Moreover, there exists a pair of points $u_0, v_0\in\mathrm{T}$ with maximal distance. Namely,

$$d(u_0,v_0) = \sup_{u,v\in\mathrm{T}} d(u,v) = D . \tag{2.30}$$

Without loss of generality, we assume that $d(u_0,\rho)\geq d(v_0,\rho)$. Then the total height of T is attained at u_0. Namely

$$d(u_0,\rho) = \sup_{u\in T} d(u,\rho) = \Gamma . \tag{2.31}$$

Proof. Let $u,v\in\mathrm{T}$. Recall from the definition of real trees (given in Introduction) that $[\![u,v]\!]$ stands for the unique geodesic path between u and v. To simplify notation, we set $h(u) := d(u,\rho)$ for $u\in\mathrm{T}$. The *branch point* $u\wedge v$ of u and v is the unique point of T satisfying

$$[\![\rho, u\wedge v]\!] = [\![\rho,u]\!]\cap[\![\rho,v]\!] .$$

Then, we easily check

$$d(u,v) = d(u,u\wedge v) + d(u\wedge v, v) = h(u) + h(v) - 2h(u\wedge v) .$$

The triangle inequality easily implies that $D \leq 2\Gamma$ while the inequality $\Gamma \leq D$ is a consequence of the definitions of Γ and D. As $d : \mathrm{T}^2 \to \mathbb{R}_+$ is continuous and T is compact, there exists a pair of points $u_0, v_0 \in \mathrm{T}$ such that (2.30) holds true. To prove (2.31), we argue by contradiction: we assume that there exists $w \in \mathrm{T}$ such that $h(w) > h(u_0)$. Let us write $b := u_0 \wedge v_0$. Here we enumerate the three possible locations of w. See Figure 2.1.

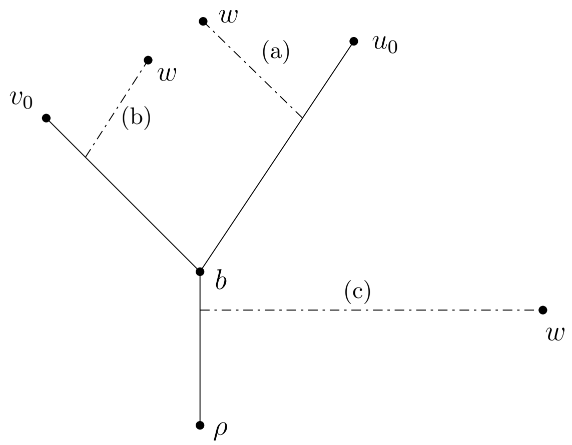

***Figure 2.1** – Three possibilities for w*

(a) Suppose that $w \wedge u_0 \in [\![u_0, b]\!]$. By hypothesis, we have $h(w) > h(u_0)$. In other words,

$$h(w) = d(w, b) + h(b) > h(u_0) = d(u_0, b) + h(b).$$

Thus, $d(w, b) > d(u_0, b)$ and

$$d(w, v_0) = d(w, b) + d(b, v_0) > d(u_0, b) + d(b, v_0) = d(u_0, v_0),$$

which contradicts (2.30).

(b) Suppose that $w \wedge v_0 \in [\![v_0, b]\!]$. In this case, we have

$$h(w) = d(w, b) + h(b) > h(u_0) \geq h(v_0) = d(v_0, b) + h(b).$$

Then $d(w, b) > d(v_0, b)$ and

$$d(w, u_0) = d(w, b) + d(b, u_0) > d(v_0, b) + d(b, u_0) = d(u_0, v_0).$$

This again contradicts (2.30).

(c) Suppose that $w \wedge u_0 \in [\![\rho, b]\!]$. Then we deduce from

$$h(w) = d(w, w \wedge u_0) + h(w \wedge u_0) > h(u_0) = d(u_0, w \wedge u_0) + h(w \wedge u_0)$$

that $d(w, w \wedge u_0) > d(u_0, w \wedge u_0)$. Note that in this case $w \wedge u_0 = w \wedge v_0$. Therefore,

$$\begin{aligned} d(w, v_0) = d(w, w \wedge v_0) + d(w \wedge v_0, v_0) &= d(w, w \wedge u_0) + d(w \wedge u_0, v_0) \\ &> d(u_0, w \wedge u_0) + d(w \wedge u_0, v_0) \\ &> d(u_0, b) + d(b, v_0) = d(u_0, v_0), \end{aligned}$$

which contradicts (2.30).

In brief, there exists no $w \in \mathrm{T}$ such that $h(w) = d(w, \rho) > h(u_0) = d(u_0, \rho)$, which entails (2.31). ■

Williams' decomposition of Brownian excursions. Let us recall the classical result of Williams' path decomposition of Brownian excursions (see for instance Revuz & Yor [97] Chaper XII Theorem 4.5). Define

$$\tau_* := \inf\{t>0 : e_t = \max_{s\geq 0} e_s\} \; . \tag{2.32}$$

Under $\mathbf{N}$ (and also under $\mathbf{N}_{\mathrm{nr}}$), τ_* is the unique time at which e reaches its maximum value. Recall from (2.11) the definition of the total height Γ of the Brownian tree coded by e. Then, we have $\Gamma = e_{\tau_*}$.

We also recall the distribution of Γ under $\mathbf{N}$:

$$\mathbf{N}\big(\Gamma \in dr\big) = \frac{dr}{r^2} \; . \tag{2.33}$$

See Revuz & Yor [97] Chaper XII Theorem 4.5 combined with Remark 16.

Williams's decomposition entails that there is a regular version of the family of conditioned laws $\mathbf{N}(\,\cdot\,|\,\Gamma = r)$, $r > 0$. Namely, $\mathbf{N}(\,\cdot\,|\,\Gamma = r)$-a.s. $\Gamma = r$, $r \mapsto \mathbf{N}(\,\cdot\,|\,\Gamma = r)$ is weakly continuous on $\mathbb{C}(\mathbb{R}_+, \mathbb{R}_+)$ and

$$\mathbf{N} = \int_0^\infty \mathbf{N}(\Gamma \in dr)\, \mathbf{N}(\,\cdot\,|\,\Gamma = r) \; . \tag{2.34}$$

Let $Z = (Z_t)_{t\geq 0}$ be a continuous process defined on the probability space $(\Omega, \mathcal{F}, \mathbf{P})$ such that $\frac{1}{\sqrt{2}} Z$ is distributed as a Bessel process of dimension 3 starting from 0. Let $\tau_r = \inf\{t>0 : Z_t = r\}$ be the hitting time of Z at level $r\in (0,\infty)$. We recall that

$$\forall \lambda \in \mathbb{R}_+, \quad \mathbf{E}\big[e^{-\lambda \tau_r}\big] = \frac{r\sqrt{\lambda}}{\sinh(r\sqrt{\lambda})} \; . \tag{2.35}$$

See Borodin & Salminen [35] Part II, Chapter 5, Section 2, Formula 2.0.1, p. 463, where we let x tend to 0 and take $\alpha = \lambda$ and $z = r/\sqrt{2}$, since $Z = \sqrt{2} R^{(3)}$.

We next introduce the following notation:

$$\overleftarrow{e}(t) = e_{(\tau_* - t)_+}; \quad \overrightarrow{e}(t) = e_{\tau_* + t}, \quad t \geq 0.$$

where $(\cdot)_+$ stands for the positive part function. *Williams' decomposition* of Brownian excursion asserts the following.

> *For all $r\in(0,\infty)$, under $\mathbf{N}(\,\cdot\,|\,\Gamma = r)$, the two processes $\overleftarrow{e}$ and $\overrightarrow{e}$ are distributed as two independent copies of $(Z_{(\tau_r - t)_+})_{t\geq 0}$.*

As a combined consequence of this decomposition and (2.35), we have

$$\forall r\in(0,\infty), \quad \mathbf{N}\big(e^{-\lambda\zeta}\,|\,\Gamma = r\big) = \mathbf{E}\big[e^{-\lambda\tau_r}\big]^2 = \left(\frac{r\sqrt{\lambda}}{\sinh(r\sqrt{\lambda})}\right)^2 , \tag{2.36}$$

where we recall that ζ stands for the lifetime of the excursion. Therefore,

$$\mathbf{N}\left(e^{-\lambda\zeta}\mathbf{1}_{\{\Gamma > a\}}\right) = \int_a^\infty \mathbf{N}(e^{-\lambda\zeta}\,|\,\Gamma = r)\mathbf{N}(\Gamma\in dr) = \int_a^\infty \frac{\lambda dr}{\sinh^2(r\sqrt{\lambda})} = \sqrt{\lambda}\coth(a\sqrt{\lambda}) - \sqrt{\lambda} ,$$

by (2.33) and (2.36). Combined with the fact that $\mathbf{N}(1 - e^{-\lambda\zeta}) = \sqrt{\lambda}$, this entails that

$$\mathbf{N}\Big(1 - e^{-\lambda\zeta}\mathbf{1}_{\{\Gamma\leq a\}}\Big) = \sqrt{\lambda}\coth(a\sqrt{\lambda}). \tag{2.37}$$

This equation is used in the proof of Theorem 2.1.

Spinal decomposition Let us interpret Williams' decomposition in terms of a Poisson decomposition of the Brownian excursion. To that end, we need some notation. Let $h \in \mathbb{C}(\mathbb{R}_+, \mathbb{R}_+)$ have compact support. We assume that $h(0) > 0$. For all $s \in \mathbb{R}_+$, we set $\underline{h}(s) = \inf_{0 \le u \le s} h(u)$. Let (l_i, r_i), $i \in \mathcal{I}(h)$ be the excursion intervals of $h - \underline{h}$ away from 0; namely, they are the connected components of the open set $\{s \ge 0 : h(s) - \underline{h}(s) > 0\}$. For all $i \in \mathcal{I}(h)$, we next set

$$h^i(s) = \big(h - \underline{h}\big)\big((l_i + s) \wedge r_i\big), \quad s \ge 0,$$

which is the excursion of $h - \underline{h}$ corresponding to the interval (l_i, r_i). Then we set

$$\mathcal{P}(h) = \sum_{i \in \mathcal{I}(h)} \delta_{(h(0) - h(l_i),\, h^i)},$$

that is a point measure on $\mathbb{R}_+ \times \mathbb{C}(\mathbb{R}_+, \mathbb{R}_+)$. We define

$$\mathcal{Q} := \mathcal{P}(\overrightarrow{e}) + \mathcal{P}(\overleftarrow{e}) =: \sum_{j \in \mathcal{J}} \delta_{(s_j, e^j)} \ . \tag{2.38}$$

We also introduce for all $t \in (0, \infty)$ the following notation:

$$\mathbf{N}_t = \mathbf{N}\big(\, \cdot \cap \{\Gamma \le t\}\big) \ . \tag{2.39}$$

The following lemma is the special case of a general result due to Abraham & Delmas [3].

Lemma 2.4 (Proposition 1.1, Abraham & Delmas [3]). *Let $r \in (0, \infty)$. Then, $\mathcal{Q}$ under $\mathbf{N}(\,\cdot \,|\, \Gamma = r)$ is a Poisson point measure on $\mathbb{R}_+ \times \mathbb{C}(\mathbb{R}_+, \mathbb{R}_+)$ with intensity measure $2 \cdot \mathbf{1}_{[0,r]}(t) dt\, \mathbf{N}_t$.*

Interpretation in terms of the Brownian tree and consequences. Let us interpret $\mathcal{Q}$ in terms of the Brownian tree $\mathcal{T}$ coded by the Brownian excursion e. Recall that $p : [0, \zeta] \to \mathcal{T}$ stands for the canonical projection and recall that $\rho = p(0)$ is the root of $\mathcal{T}$. The point $p(\tau_*)$ is the (unique) point of $\mathcal{T}$ that attains the total height: $d(\rho, p(\tau_*)) = \Gamma$.

Denote by $\mathcal{T}^o_{j'}$, $j' \in \mathcal{J}'$, the connected components of $\mathcal{T} \backslash [\![\rho, p(\tau_*)]\!]$. For all $j' \in \mathcal{J}'$, there exists a unique point $\sigma_{j'} \in [\![\rho, p(\tau_*)]\!]$ such that $\mathcal{T}_{j'} := \mathcal{T}^o_{j'} \cup \{\sigma_{j'}\}$ is the closure of $\mathcal{T}^o_{j'}$ in $\mathcal{T}$. Recall the notation $\mathcal{J}$ from (2.38). It is not difficult to see that $\mathcal{J}'$ is in one-to-one correspondence with $\mathcal{J}$. Moreover, after a re-indexing, we can suppose that $d(p(\tau_*), \sigma_j) = s_j$ and that $(\mathcal{T}_j, d, \sigma_j)$ is the real tree coded by the excursion e^j, for each $j \in \mathcal{J}$. Then we set

$$\forall j \in \mathcal{J}, \quad \Gamma_j := \max_{s \ge 0} e^j(s) = \max_{\gamma \in \mathcal{T}_j} d(\sigma_j, \gamma) \ , \tag{2.40}$$

that is the total height of the rooted real tree $(\mathcal{T}_j, d, \sigma_j)$. We claim that

$$\mathbf{N}\text{-a.e.} \quad D = \sup_{j \in \mathcal{J}} (s_j + \Gamma_j) \ . \tag{2.41}$$

Proof of (2.41). First observe that for all $t \in (0, \infty)$, $\mathbf{N}_t$ is an infinite measure because $\mathbf{N}$ is infinite and because $\mathbf{N}(\Gamma > t) = 1/t$ by (2.33). By Lemma 2.4, $\mathbf{N}$-a.e. the closure of the set $\{s_j \,;\, j \in \mathcal{J}\}$ is $[0, \Gamma]$. This entails that

$$\mathbf{N}\text{-a.e.} \quad \Gamma = \sup_{j \in \mathcal{J}} s_j \le \sup_{j \in \mathcal{J}} (s_j + \Gamma_j) \ . \tag{2.42}$$

Next, for all $j \in \mathcal{J}$, there exists $\gamma_j \in \mathcal{T}_j$ such that $d(\sigma_j, \gamma_j) = \Gamma_j$. Then observe that

$$d(p(\tau_*), \gamma_j) = d(p(\tau_*), \sigma_j) + d(\sigma_j, \gamma_j) = s_j + \Gamma_j \ . \tag{2.43}$$

Note that Lemma 2.3 implies that $D = \max_{\gamma\in\mathcal{T}} d(p(\tau_*), \gamma)$. Comparing this with (2.43), we get

$$D \geq \sup_{j\in\mathcal{J}} (s_j + \Gamma_j) \; . \tag{2.44}$$

On the other hand, there exists $\gamma^* \in \mathcal{T}$ such that $D = \max_{\gamma\in\mathcal{T}} d(p(\tau_*), \gamma) = d(p(\tau_*), \gamma^*)$ by Lemma 2.3. If $\gamma^* \notin [\![\rho, p(\tau_*)]\!]$, then there exists $j^* \in \mathcal{J}$ such that $\gamma^* \in \mathcal{T}_{j^*}$. In consequence, we have $D = d(p(\tau_*), \gamma^*) \leq s_{j^*} + \Gamma_{j^*}$, and then $D = \sup_{j\in\mathcal{J}} (s_j + \Gamma_j)$ when compared with (2.44). If $\gamma^* \in [\![\rho, p(\tau_*)]\!]$, then (2.42) implies that $\gamma^* = \rho$ and $D = \Gamma$. In both cases (2.41) holds true. ■

We next denote by ζ_j the lifetime of e^j for all $j \in \mathcal{J}$ and prove the following statement.

$$\mathbf{N}\text{-a.e.} \quad \sum_{j\in\mathcal{J}} \zeta_j = \zeta \; . \tag{2.45}$$

Proof of (2.45). Let $\sigma \in [\![p(\tau_*), \rho]\!]$ be distinct from $p(\tau_*)$ and ρ; then $\mathrm{n}(\sigma) \geq 2$ and σ is not a leaf of $\mathcal{T}$. Recall from (2.12) the definition of the mass measure m and recall from (2.13) that $\mathbf{N}_{\mathrm{nr}}$-a.s. m is diffuse and supported on the set of leaves of $\mathcal{T}$. By (2.7), this property also holds true $\mathbf{N}$-almost everywhere and we thus get

$$\mathbf{N}\text{-a.e.} \quad \mathrm{m}\big([\![p(\tau_*), \rho]\!]\big) = 0 \; .$$

Recall that $\mathcal{T}_j^o$, $j \in \mathcal{J}$, are the connected components of $\mathcal{T} \backslash [\![\rho, p(\tau_*)]\!]$. Thus,

$$\mathbf{N}\text{-a.e.} \quad \mathrm{m}(\mathcal{T}) = \mathrm{m}\big([\![p(\tau_*), \rho]\!]\big) + \sum_{j\in\mathcal{J}} \mathrm{m}\big(\mathcal{T}_j^o\big) = \sum_{j\in\mathcal{J}} \mathrm{m}\big(\mathcal{T}_j^o\big). \tag{2.46}$$

Recall that $\mathcal{T}_j = \mathcal{T}_j^o \cup \{\sigma_j\}$ and that m is $\mathbf{N}$-a.e. diffuse, which entails $\mathrm{m}(\mathcal{T}_j) = \mathrm{m}(\mathcal{T}_j^o)$, for all $j \in \mathcal{J}$. Moreover, since $(\mathcal{T}_j, d, \sigma_j)$ is coded by the excursion e^j, we have $\zeta_j = \mathrm{m}(\mathcal{T}_j)$. For a similar reason, we also have $\zeta = \mathrm{m}(\mathcal{T})$. This, combined with (2.46), entails (2.45). ■

2.3 Proof of Theorem 2.1

First we note that by (2.9),

$$\mathrm{L}_\lambda(y, z) = \frac{1}{2\sqrt{\pi}} \int_0^\infty dr e^{-\lambda r} r^{-\frac{3}{2}} \mathbf{N}_{\mathrm{nr}}\big(r^{\frac{1}{2}} D > 2y \, ; \, r^{\frac{1}{2}} \Gamma > z\big) = \mathbf{N}\left(e^{-\lambda\zeta} \mathbf{1}_{\{D>2y;\Gamma>z\}}\right). \tag{2.47}$$

Observe that the scaling property (2.16) is a direct consequence of the scaling property of $\mathbf{N}$ (see (2.6)).

We next compute the right hand side of (2.47). To that end, recall from (2.38) the spinal decomposition of the excursion e and recall from (2.40) the notation $\Gamma_j = \max_{s\geq 0} e^j(s)$, for all $j \in \mathcal{J}$; also recall that ζ_j stands for the lifetime of e^j. Let $r, y \in (0, \infty)$ be such that $y \leq r \leq 2y$. We apply successively (2.41), (2.45), Lemma 2.4 and Campbell's formula for Poisson point measures to get

$$\begin{aligned} \mathbf{N}\left(e^{-\lambda\zeta} \mathbf{1}_{\{D\leq 2y\}} \,\Big|\, \Gamma = r\right) &= \mathbf{N}\Big(\prod_{j\in\mathcal{J}} e^{-\lambda\zeta_j} \mathbf{1}_{\{s_j+\Gamma_j \leq 2y\}} \,\Big|\, \Gamma = r\Big) \\ &= \exp\left(-2\int_0^r dt \, \mathbf{N}_t\big(1 - e^{-\lambda\zeta} \mathbf{1}_{\{\Gamma\leq 2y-t\}}\big)\right) \; . \end{aligned} \tag{2.48}$$

Recall from (2.39) that $\mathbf{N}_t = \mathbf{N}\big(\,\cdot\, \cap \{\Gamma \leq t\}\big)$ and observe that

$$\int_0^r dt \, \mathbf{N}_t\big(1 - e^{-\lambda\zeta} \mathbf{1}_{\{\Gamma\leq 2y-t\}}\big) = \int_0^y dt \, \mathbf{N}\big((1 - e^{-\lambda\zeta}) \mathbf{1}_{\{\Gamma\leq t\}}\big) + \int_y^r dt \, \mathbf{N}\big(\mathbf{1}_{\{\Gamma\leq t\}} - e^{-\lambda\zeta} \mathbf{1}_{\{\Gamma<2y-t\}}\big) \; . \tag{2.49}$$

By (2.37) and by (2.33),

$$\mathbf{N}\big((1-e^{-\lambda\zeta})\mathbf{1}_{\{\Gamma\le t\}}\big) = \mathbf{N}\big(1-e^{-\lambda\zeta}\mathbf{1}_{\{\Gamma\le t\}}\big) - \mathbf{N}\big(\Gamma > t\big) = \sqrt{\lambda}\coth\big(t\sqrt{\lambda}\big) - \tfrac{1}{t} \tag{2.50}$$

and

$$\mathbf{N}\big(\mathbf{1}_{\{\Gamma\le t\}} - e^{-\lambda\zeta}\mathbf{1}_{\{\Gamma<2y-t\}}\big) = \mathbf{N}\big(1-e^{-\lambda\zeta}\mathbf{1}_{\{\Gamma\le 2y-t\}}\big) - \mathbf{N}(\Gamma > t) = \sqrt{\lambda}\coth\big((2y-t)\sqrt{\lambda}\big) - \tfrac{1}{t}. \tag{2.51}$$

Then observe that for all $\varepsilon, a \in (0,\infty)$ such that $\varepsilon < a$,

$$\int_\varepsilon^a \big(\sqrt{\lambda}\coth(t\sqrt{\lambda}) - \tfrac{1}{t}\big)dt = \log\frac{\sinh a\sqrt{\lambda}}{a} - \log\frac{\sinh\varepsilon\sqrt{\lambda}}{\varepsilon}.$$

Thus, as $\varepsilon \to 0$, we get

$$\forall a \in \mathbb{R}_+ \quad \int_0^a \big(\sqrt{\lambda}\coth(t\sqrt{\lambda}) - \tfrac{1}{t}\big)dt = \log\frac{\sinh a\sqrt{\lambda}}{a\sqrt{\lambda}}\,. \tag{2.52}$$

An easy computation based on (2.52) and combined with (2.49), (2.50), (2.51) and (2.48) entails

$$\mathbf{N}\Big(e^{-\lambda\zeta}\mathbf{1}_{\{D\le 2y\}}\Big|\,\Gamma = r\Big) = \frac{\Big(r\sqrt{\lambda}\sinh\big((2y-r)\sqrt{\lambda}\big)\Big)^2}{\sinh^4(y\sqrt{\lambda})}.$$

Combining this with (2.36), we get

$$\forall r, y \in (0,\infty) : y \le r \le 2y,\ \mathbf{N}\big(e^{-\lambda\zeta}\mathbf{1}_{\{D>2y\}}\big|\Gamma = r\big) = \left(\frac{r\sqrt{\lambda}}{\sinh(r\sqrt{\lambda})}\right)^2 - \frac{\Big(r\sqrt{\lambda}\sinh\big((2y-r)\sqrt{\lambda}\big)\Big)^2}{\sinh^4(y\sqrt{\lambda})}. \tag{2.53}$$

Next, let $r, y \in (0,\infty)$ be such that $r > 2y$. By Lemma 2.3, $\Gamma \le D \le 2\Gamma$. Therefore,

$$\forall r, y \in (0,\infty) : r > 2y,\ \mathbf{N}\Big(e^{-\lambda\zeta}\mathbf{1}_{\{D>2y\}}\Big|\,\Gamma = r\Big) = \mathbf{N}\Big(e^{-\lambda\zeta}\Big|\,\Gamma = r\Big) = \left(\frac{r\sqrt{\lambda}}{\sinh(r\sqrt{\lambda})}\right)^2. \tag{2.54}$$

Finally, let $r < y$. Then $\mathbf{N}(e^{-\lambda\zeta}\mathbf{1}_{\{D>2y\}}|\Gamma = r) = 0$, since $\Gamma \le D \le 2\Gamma$. Combining this with (2.53) and (2.54), we easily obtain that

$$\begin{aligned}
\mathbf{N}\left(e^{-\lambda\zeta}\mathbf{1}_{\{D>2y,\Gamma>z\}}\right) &= \int_z^\infty \mathbf{N}\Big(e^{-\lambda\zeta}\mathbf{1}_{\{D>2y\}}\Big|\,\Gamma = r\Big)\mathbf{N}(\Gamma \in dr) \\
&= \int_{z\vee y}^{2y\wedge z} \mathbf{N}\Big(e^{-\lambda\zeta}\mathbf{1}_{\{D>2y\}}\Big|\,\Gamma = r\Big)\mathbf{N}(\Gamma \in dr) + \int_{2y}^\infty \mathbf{N}\Big(e^{-\lambda\zeta}\Big|\,\Gamma = r\Big)\mathbf{N}(\Gamma \in dr) \\
&= \sqrt{\lambda}\big(\coth\big((z\vee y)\sqrt{\lambda}\big) - 1\big) - \mathbf{1}_{\{z\le 2y\}}\frac{\sqrt{\lambda}\sinh(2q\sqrt{\lambda}) - 2\lambda q}{4\sinh^4(y\sqrt{\lambda})}\,,
\end{aligned}$$

where we recall the notation $q = y\wedge(2y-z)$. By (2.47), this concludes the proof of Theorem 2.1.

2.4 Proof of Corollary 2.2

We introduce the following notation for the Laplace transform on $\mathbb{R}_+$: for all Lebesgue integrable functions $f : \mathbb{R}_+ \to \mathbb{R}$, we set

$$\forall \lambda \in \mathbb{R}_+, \qquad \mathcal{L}_\lambda(f) := \int_0^\infty dx\, e^{-\lambda x} f(x),$$

which is well-defined. Note that if f, g are two continuous and integrable functions such that $\mathcal{L}_\lambda(f) = \mathcal{L}_\lambda(g)$ for all $\lambda \in [0, \infty)$, then we have $f = g$, by the injectiveness of Laplace transform and standard arguments.

For all $a, x \in (0, \infty)$, we set $f_a(x) = \frac{a}{2\sqrt{\pi}} x^{-3/2} e^{-a^2/4x}$. It is well-known that $\mathcal{L}_\lambda(f_a) = e^{-a\sqrt{\lambda}}$ for all $\lambda \in \mathbb{R}_+$ (see for instance Borodin & Salminen [35] Appendix 3, Particular formulæ 2, p. 650). Then we set

$$g_a(x) = \partial_x f_a(x) = \frac{1}{8\sqrt{\pi}} x^{-\frac{7}{2}} e^{-\frac{a^2}{4x}} (a^3 - 6ax) \quad \text{and} \quad h_a(x) = -\partial_a f_a(x) = \frac{1}{4\sqrt{\pi}} x^{-\frac{5}{2}} e^{-\frac{a^2}{4x}} (a^2 - 2x) \; .$$

Consequently, for all $\lambda \in \mathbb{R}_+$,

$$\mathcal{L}_\lambda(g_a) = \lambda e^{-a\sqrt{\lambda}} \quad \text{and} \quad \mathcal{L}_\lambda(h_a) = \sqrt{\lambda} e^{-a\sqrt{\lambda}} \; . \tag{2.55}$$

(See also Borodin & Salminen [35] Appendix 3, Particular formulæ 3 and 4, p. 650.) Moreover, we have the following easy bounds: for all $\lambda \in \mathbb{R}_+$,

$$\mathcal{L}_\lambda(|g_a|) \leq \frac{1}{8\sqrt{\pi}} \int_0^\infty dx \, e^{-\lambda x} x^{-\frac{7}{2}} e^{-\frac{a^2}{4x}} (a^3 + 6ax) = \lambda e^{-a\sqrt{\lambda}} + \tfrac{6}{a} \sqrt{\lambda} e^{-a\sqrt{\lambda}} + \tfrac{6}{a^2} e^{-a\sqrt{\lambda}}, \tag{2.56}$$

$$\mathcal{L}_\lambda(|h_a|) \leq \frac{1}{4\sqrt{\pi}} \int_0^\infty dx \, e^{-\lambda x} x^{-\frac{5}{2}} e^{-\frac{a^2}{4x}} (a^2 + 2x) = \sqrt{\lambda} e^{-a\sqrt{\lambda}} + \tfrac{2}{a} e^{-a\sqrt{\lambda}}. \tag{2.57}$$

Let $y, z \in (0, \infty)$. Recall from (2.20) the notation ρ and δ. Next set

$$\forall n \in \mathbb{N}, \quad u_n = \tfrac{1}{6}(n+3)(n+2)(n+1),$$

so that $(1-x)^{-4} = \sum_{n\geq 0} u_n x^n$, for all $x \in [0, 1)$. Then (2.17) implies that

$$\begin{aligned} \mathrm{L}_1(\tfrac{1}{2}y, z) &= \coth \rho - 1 - \frac{\sinh(\delta y) - \delta y}{4 \sinh^4(y/2)} = \frac{2e^{-2\rho}}{1 - e^{-2\rho}} + \frac{2e^{-2y}(e^{-\delta y} - e^{\delta y})}{(1 - e^{-y})^4} + \frac{4\delta y e^{-2y}}{(1 - e^{-y})^4} \\ &= \sum_{n\geq 1} 2e^{-2n\rho} + \sum_{n\geq 0} 2u_n \big(e^{-(n+2+\delta)y} - e^{-(n+2-\delta)y} + 2\delta y e^{-(n+2)y} \big) \\ &= \sum_{n\geq 1} 2e^{-2n\rho} + \sum_{n\geq 2} 2u_{n-2} \big(e^{-(n+\delta)y} - e^{-(n-\delta)y} + 2\delta y e^{-ny} \big) \; . \end{aligned}$$

Thus, by (2.16), we obtain that

$$\begin{aligned} \frac{1}{2\sqrt{\pi}} \int_0^\infty e^{-\lambda r} r^{-\frac{3}{2}} \, \mathbf{N}_{\mathrm{nr}} \big(r^{\frac{1}{2}} D > y \, ; \, r^{\frac{1}{2}} \Gamma > z \big) \, dr &= \mathrm{L}_\lambda(\tfrac{1}{2}y, z) = \sqrt{\lambda} \mathrm{L}_1(\tfrac{1}{2}y\sqrt{\lambda}, z\sqrt{\lambda}) \\ &= \sum_{n\geq 1} 2\sqrt{\lambda} e^{-2n\rho\sqrt{\lambda}} + \sum_{n\geq 2} 2u_{n-2} \big(\sqrt{\lambda} e^{-(n+\delta)y\sqrt{\lambda}} - \sqrt{\lambda} e^{-(n-\delta)y\sqrt{\lambda}} + 2\delta y \lambda e^{-ny\sqrt{\lambda}} \big) \\ &= \sum_{n\geq 1} 2\mathcal{L}_\lambda(h_{2n\rho}) + \sum_{n\geq 2} 2u_{n-2} \mathcal{L}_\lambda \big(h_{(n+\delta)y} - h_{(n-\delta)y} + 2\delta y g_{ny} \big) \; . \end{aligned} \tag{2.58}$$

Observe that for all $r \in \mathbb{R}_+$,

$$\sum_{n\geq 1} 2 \sup_{[0,r]} |h_{2n\rho}| + \sum_{n\geq 2} 2u_{n-2} \big(\sup_{[0,r]} |h_{(n+\delta)y}| + \sup_{[0,r]} |h_{(n-\delta)y}| + 2\delta y \sup_{[0,r]} |g_{ny}| \big) < \infty \; . \tag{2.59}$$

Then, for any $r \in \mathbb{R}_+$, we set

$$\phi_{y,z}(r) := 2 \sum_{n=1}^{\infty} h_{2n\rho}(r) e^{-r} + \sum_{n=2}^{\infty} 2u_{n-2} \big(h_{(n+\delta)y}(r) e^{-r} - h_{(n-\delta)y}(r) e^{-r} + 2\delta y g_{ny}(r) e^{-r} \big) \; ,$$

which is well-defined and continuous thanks to (2.59). The bounds (2.56) and (2.57) imply that $\phi_{y,z}$ is Lebesgue integrable. Moreover, (2.58) asserts that $\mathrm{L}_{\lambda+1}(\frac{1}{2}y, z) = \mathcal{L}_\lambda(\phi_{y,z})$. By the injectiveness of Laplace transform for continuous integrable functions (as mentioned above), we get

$$\forall r \in \mathbb{R}_+, \quad \phi_{y,z}(r) = \frac{1}{2\sqrt{\pi}} e^{-r} r^{-\frac{3}{2}} \, \mathbf{N}_{\mathrm{nr}}\big(r^{\frac{1}{2}} D > y \, ; \, r^{\frac{1}{2}} \Gamma > z \big) \; ,$$

which entails (2.21) by taking $r = 1$.

Since $\Gamma \le D \le 2\Gamma$, if $z = y$, then $\mathbf{N}_{\mathrm{nr}}(D > y; \Gamma > y) = \mathbf{N}_{\mathrm{nr}}(\Gamma > y)$ and (2.21) immediately implies (2.23) because in this case $\rho = y$ and $\delta = 0$. If $z = \frac{1}{2}y$, then $\mathbf{N}_{\mathrm{nr}}(D > y; \Gamma > \frac{1}{2}y) = \mathbf{N}_{\mathrm{nr}}(D > y)$, $\rho = \frac{1}{2}y$, $\delta = 1$ and (2.21) implies

$$\mathbf{N}_{\mathrm{nr}}(D > y) = \sum_{n\ge 1} \big(n^2y^2 - 2\big)\, e^{-\frac{1}{4}n^2y^2} + \frac{1}{6}\sum_{n\ge 2} n(n^2-1)\Big[\big[((n+1)y)^2 - 2\big] e^{-\frac{1}{4}((n+1)y)^2} - \\ \big[((n-1)y)^2 - 2\big] e^{-\frac{1}{4}((n-1)y)^2} + y(n^3y^3 - 6ny)e^{-\frac{1}{4}n^2y^2} \Big],$$

which entails (2.24) by re-indexing the sums according to $e^{-n^2y^2/4}$: we leave the details to the reader. We next derive (2.27) by differentiating (2.24). As mentioned in Remark 18, we use Jacobi identity (2.29) to derive (2.22) from (2.21). The computations are long but straightforward: we leave them to the reader. Finally, for the same reason as before, (2.22) entails (2.25) by taking $\rho = y$ and $\delta = 0$. It also entails (2.26) by taking $\rho = \frac{1}{2}y$ and $\delta = 1$. Differentiating (2.26) gives (2.28). This completes the proof of Corollary 2.2.

Chapter 3

Decomposition of Lévy trees along their diameter

The results of this chapter are from the joint work [54] with Thomas Duquesne, submitted for publication.

Contents

We study the diameter of Lévy trees that are random compact metric spaces obtained as the scaling limits of Galton-Watson trees. Lévy trees have been introduced by Le Gall and Le Jan (1998) and they generalise Aldous' Continuum Random Tree (1991) that corresponds to the Brownian case. We first characterize the law of the diameter of Lévy trees and we prove that it is realized by a unique pair of points. We prove that the law of Lévy trees conditioned to have a fixed diameter $r \in (0, \infty)$ is obtained by glueing at their respective roots two independent size-biased Lévy trees conditioned to have height $r/2$ and then by uniformly re-rooting the resulting tree; we also describe by a Poisson point measure the law of the subtrees that are grafted on the diameter. As an application of this decomposition of Lévy trees according to their diameter, we characterize the joint law of the height and the diameter of stable Lévy trees conditioned by their total mass; we also provide asymptotic expansions of the law of the height and of the diameter of such normalized stable trees, which generalizes the identity due to Szekeres (1983) in the Brownian case.

3.1 Introduction and main results

Lévy tree are random compact metric spaces that are the scaling limits of Galton-Watson trees. The Brownian tree, also called the continuum random tree, is a particular instance of Lévy trees; it is the limit of the rescaled uniformly distributed rooted labelled tree with n vertices. The Brownian tree has been introduced by Aldous in [8] and further studied in Aldous [9, 10]. Lévy trees have been introduced by Le Gall & Le Jan [83] via a coding function called the height process that is a local time functional of a spectrally positive Lévy process. Lévy trees (and especially stable trees) have been studied in D. & Le Gall [51, 52] (geometric and fractal properties, connection with superprocesses), see D. & Winkel [55] and Marchal [87] for alternative constructions, see also Miermont [89, 90], Haas & Miermont [65], Goldschmidt & Haas [61] for applications to stable fragmentations, and Abraham & Delmas [1, 2], Abraham, Delmas & Voisin [6] for general fragmentations and pruning processes on Lévy trees.

In this article, we study the diameter of Lévy trees. As observed by Aldous (see [9], Section 3.4), in the Browian case the law of the diameter has been found by Szekeres [98] by taking the limit of the generating function of the diameter of uniformly distributed rooted labelled tree with n vertices. Then, the question was raised by Aldous that whether we can derive the law of the diameter directly from the normalised Brownian excursion that codes the Brownian tree (see also Pitman [95], Exercise 9.4.1). This question is now answered in W. [100].

In this article we compute the law of the diameter for general Lévy trees (see Theorem 3.1). We also prove that the diameter of Lévy trees is realized by a unique pair of points. The geodesic path joining these two extremal points is therefore unique. In Theorem 3.2, we describe the coding function (the height process) of the Lévy trees rerooted at the midpoint of their diameter, which plays the role of an intrinsic root. The proof of Theorem 3.2 that provides a decomposition of Lévy trees according to their diameter specifically relies on the invariance of Lévy trees by uniform rerooting, as proved by D. & Le Gall in [53], and on the decomposition of Lévy trees according to their height, as proved by Abraham & Delmas [3] (this decomposition generalizes Williams' decomposition of the Brownian excursion). Roughly speaking, Theorem 3.2 asserts that a Lévy tree that is conditioned to have diameter r and that is rooted at its midpoint is obtained by glueing at their root two size-biased independent Lévy trees conditioned to have height $r/2$ and then by rerooting uniformly the resulting tree; Theorem 3.2 also explains the distribution of the trees grafted on the diameter. As an application of this theorem, we characterize the joint law of the height and the diameter of stable trees conditioned by their total mass (see Proposition 3.3) and by providing an asymptotic expansion of the law of the height (Theorem 3.5) and of the law of the diameter (Theorem 3.7). These two asymptotic expansions generalize the identities of Szekeres in the Brownian case which involves theta functions (these identities are recalled in (3.50) and (3.51)). Before stating precisely our main results we need to recall definitions and to set notations.

Real trees. We first define real-trees that are metric spaces generalizing graph-trees: let (T, d) be a metric space; it is a *real tree* iff the following holds true.

(a) For any $\sigma_1, \sigma_1 \in T$, there is a unique isometry $f : [0, d(\sigma_1, \sigma_2)] \to T$ such that $f(0) = \sigma_1$ and $f(d(\sigma_1, \sigma_2)) = \sigma_2$. Then, we shall use the following notation: $[\![\sigma_1, \sigma_2]\!] := f([0, d(\sigma_1, \sigma_2)])$.

(b) For any continuous injective function $q : [0, 1] \to T$, $q([0, 1]) = [\![q(0), q(1)]\!]$.

When a point $\rho \in T$ is distinguished, (T, d, ρ) is said to be a *rooted* real tree, ρ being the *root* of T. Among connected metric spaces, real trees are characterized by the so-called *four points inequality* that is expressed as follows: let (T, d) be a connected metric space; then (T, d) is a real tree iff for any $\sigma_1, \sigma_2, \sigma_3, \sigma_4 \in T$, we have

$$d(\sigma_1, \sigma_2) + d(\sigma_3, \sigma_4) \leq \big(d(\sigma_1, \sigma_3) + d(\sigma_2, \sigma_4)\big) \vee \big(d(\sigma_1, \sigma_4) + d(\sigma_2, \sigma_3)\big). \tag{3.1}$$

We refer to Evans [57] or to Dress, Moulton and Terhalle [47] for a detailed account on this property. Let us briefly mention that the set of (pointed) isometry classes of compact rooted real trees can be equipped with the (pointed) Gromov-Hausdorff distance that makes it a Polish space: see Evans, Pitman & Winter [59], Theorem 2, for more details on this intrinsic point of view on trees that we shall not use here.

The coding of real tree. Let us briefly recall how real trees can be obtained thanks to continuous functions. To that end we denote by $\mathbb{C}(\mathbb{R}_+, \mathbb{R}_+)$ the space of $\mathbb{R}_+$-valued continuous function equipped with the topology of the uniform convergence on every compact subsets of $\mathbb{R}_+$. We shall denote by $H = (H_t)_{t\geq 0}$ the canonical process on $\mathbb{C}(\mathbb{R}_+, \mathbb{R}_+)$. First assume that H has a compact support, that $H_0 = 0$ and that H is distinct from the null function: we call such a function a *coding function* and we then set $\zeta_H = \sup\{t > 0 : H_t > 0\}$ that is called the *lifetime* of the coding function H. Note that $\zeta_H \in (0, \infty)$. Then, for every $s, t \in [0, \zeta_H]$, we set

$$b_H(s,t) = \inf_{r\in[s\wedge t, s\vee t]} H_r \quad \text{and} \quad d_H(s,t) = H_s + H_t - 2b_H(s,t). \tag{3.2}$$

It is easy to check that d_H satisfies the four points inequality: namely, for all $s_1, s_2, s_3, s_4 \in [0, \zeta_H]$, $d_H(s_1, s_2) + d_H(s_3, s_4) \leq \big(d_H(s_1, s_3) + d_H(s_2, s_4)\big) \vee \big(d_H(s_1, s_4) + d_H(s_2, s_3)\big)$. By taking $s_3 = s_4$, we see that d_H is a pseudometric on $[0, \zeta_H]$. We then introduce the equivalence relation $s \sim_H t$ iff $d_H(s,t) = 0$ and we set

$$\mathcal{T}_H = [0, \zeta_H]/\sim_H \ . \tag{3.3}$$

Standard arguments show that d_H induces a true metric on the quotient set $\mathcal{T}_H$ that we keep denoting d_H. We denote by $p_H : [0, \zeta_H] \to \mathcal{T}_H$ the *canonical projection*. Since H is continuous, so is p_H and $(\mathcal{T}_H, d_H)$ is therefore a compact connected metric space that satisfies the four points inequality: it is a compact real tree. We next set $\rho_H = p_H(0) = p_H(\zeta_H)$ that is chosen as the *root* of $\mathcal{T}_H$.

We next define the *total height* and the *diameter* of $\mathcal{T}_H$ that are expressed in terms of d_H as follows:

$$\Gamma_H := \sup_{\sigma\in\mathcal{T}_H} d_H(\rho_H, \sigma) = \sup_{t\in[0,\zeta_H]} H_t \quad \text{and} \quad D_H := \sup_{\sigma,\sigma'\in\mathcal{T}_H} d_H(\sigma, \sigma') = \sup_{0\leq s<t\leq\zeta_H} \big(H_s + H_t - 2 \inf_{r\in[s,t]} H_r\big) \ . \tag{3.4}$$

For any $\sigma \in \mathcal{T}_H$, we denote by $\mathrm{n}(\sigma)$ the number of connected components of the open set $\mathcal{T}_H \backslash \{\sigma\}$. Note that $\mathrm{n}(\sigma)$ is possibly infinite. We call this number the *degree* of σ. We say that σ is a *branching point* if $\mathrm{n}(\sigma) \geq 3$; we say that σ is a *leaf* if $\mathrm{n}(\sigma) = 1$ and we say that σ is *simple* if $\mathrm{n}(\sigma) = 2$. We shall use the following notation for the set of branching points and the set of leaves of $\mathcal{T}_H$:

$$\mathtt{Br}(\mathcal{T}_H) := \big\{\sigma \in \mathcal{T}_H : \mathrm{n}(\sigma) \geq 3\big\} \quad \text{and} \quad \mathtt{Lf}(\mathcal{T}_H) := \big\{\sigma \in \mathcal{T}_H : \mathrm{n}(\sigma) = 1\big\} \ . \tag{3.5}$$

In addition to the metric d_H and to the root ρ_H, the coding function yields two additional useful features: first, the *mass measure* $\mathbf{m}_H$ that is the pushforward measure of the Lebesgue measure on $[0, \zeta_H]$ induced by p_H on $\mathcal{T}_H$; namely, for any Borel measurable function $f : \mathcal{T}_H \to \mathbb{R}_+$,

$$\int_{\mathcal{T}_H} f(\sigma)\, \mathbf{m}_H(d\sigma) = \int_0^{\zeta_H} f(p_H(t))\, dt \ . \tag{3.6}$$

This measure plays an important role in the study of Lévy trees (that are defined below): in a certain sense, the mass measure is the most spread out measure on $\mathcal{T}_H$. The coding H also induced a *linear order* $\leq_H$ on $\mathcal{T}_H$ that is inherited from that of $[0, \zeta_H]$: namely for any $\sigma_1, \sigma_2 \in \mathcal{T}_H$,

$$\sigma_1 \leq_H \sigma_2 \iff \inf\{t \in [0, \zeta_H] : p_H(t) = \sigma_1\} \leq \inf\{t \in [0, \zeta_H] : p_H(t) = \sigma_2\} \ . \tag{3.7}$$

Roughly speaking, the coding function H is completely characterized by $(\mathcal{T}_H, d_H, \rho_H, \mathbf{m}_H, \leq_H)$: see D. [50] for more detail about the coding of real trees by functions.

Re-rooting trees. Several statements of our article involve a re-rooting procedure at the level of the coding functions that is recalled here from D. & Le Gall [52], Lemma 2.2 (see also [53]). Let H be a coding function as defined above and recall that $\zeta_H \in (0,\infty)$. For any $t\in\mathbb{R}_+$, denote by $\bar{t}$ the unique element of $[0,\zeta_H)$ such that $t-\bar{t}$ is an integer multiple of ζ_H. Then for all $t_0\in\mathbb{R}_+$, we set

$$\forall t\in[0,\zeta_H],\quad H^{[t_0]}_t = d_H\big(\overline{t_0}, \overline{t+t_0}\big) \quad \text{and} \quad \forall t\geq \zeta_H,\quad H^{[t_0]}_t = 0\ . \tag{3.8}$$

Then observe that $\zeta_H = \zeta_{H^{[t_0]}}$ and that

$$\forall t,t'\in[0,\zeta_H],\quad d_{H^{[t_0]}}(t,t') = d_H\big(\overline{t+t_0}, \overline{t'+t_0}\big)\ . \tag{3.9}$$

Lemma 2.2 [52] asserts that there exists a unique isometry $\phi : \mathcal{T}_{H^{[t_0]}} \to \mathcal{T}_H$ such that $\phi(p_{H^{[t_0]}}(t)) = p_H\big(\overline{t+t_0}\big)$ for all $t\in[0,\zeta_H]$. This allows to *identify canonically* $\mathcal{T}_{H^{[t_0]}}$ *with the tree* $\mathcal{T}_H$ *re-rooted at* $p_H(t_0)$:

$$\big(\mathcal{T}_{H^{[t_0]}}, d_{H^{[t_0]}}, \rho_{H^{[t_0]}}\big) \equiv \big(\mathcal{T}_H, d_H, p_H(t_0)\big)\ . \tag{3.10}$$

Note that up to this identification, $\mathbf{m}_{H^{[t_0]}}$ is the same as $\mathbf{m}_H$. Roughly speaking, the linear order $\leq_{H^{[t_0]}}$ is obtained from $\leq_H$ by a cyclic shift after $p_H(t_0)$.

Spinal decomposition. The law of the Lévy tree conditioned by its diameter that is discussed below is described as a Poisson decomposition of the trees grafted along the diameter. To explain that kind of decomposition in terms of the coding function of the tree, we introduce the following definitions and notations.

Let $h\in\mathbf{C}(\mathbb{R}_+,\mathbb{R}_+)$ have compact support. Note that $h(0)>0$ possibly. We first define the excursions of h above its infimum as follows. For any $a\in[0,h(0)]$, we first set

$$\ell_a(h) := \inf\big\{t\in\mathbb{R}_+ : h(t)=h(0)-a\big\} \quad \text{and} \quad r_a(h) := \zeta_h \wedge \inf\big\{t\in(0,\infty) : h(0)-a > h(t)\big\},$$

with the convention that $\inf\emptyset=\infty$, so that $r_{h(0)}(h)=\zeta_h$. We then set

$$\forall s\in\mathbb{R}_+,\quad \mathcal{E}_s(h,a) := h\big((\ell_a(h)+s)\wedge r_a(h)\big) - h(0) + a\ .$$

See Figure 3.1. Note that $\mathcal{E}(h,a)$ is a nonnegative continuous function with compact support such that $\mathcal{E}_0(h,a)=0$. Moreover, if $\ell_a(h)=r_a(h)$, then $\mathcal{E}(h,a)=\mathbf{0}$, the *null function*.

Let H be a coding function as defined above. Let $t\in\mathbb{R}_+$. We next set

$$\forall s\in\mathbb{R}_+,\quad H^-_s = H_{(t-s)_+} \quad \text{and} \quad H^+_s = H_{t+s}\ .$$

Note that $H^-_0=H^+_0=H_t$. To simplify notation we also set

$$\forall a\in[0,H_t],\quad \overleftarrow{H}^a := \mathcal{E}(H^-,a) \quad \text{and} \quad \overrightarrow{H}^a := \mathcal{E}(H^+,a)$$

and

$$\mathcal{J}_{0,t} := \big\{a\in[0,H_t] : \text{either } \ell_a(H^-)<r_a(H^-) \text{ or } \ell_a(H^+)<r_a(H^+)\big\}$$

that is countable. We then define the following point measure on $[0,H_t]\times\mathbf{C}(\mathbb{R}_+,\mathbb{R}_+)^2$:

$$\mathcal{M}_{0,t}(H) = \sum_{a\in\mathcal{J}_{0,t}} \delta_{(a,\overleftarrow{H}^a,\overrightarrow{H}^a)}\ , \tag{3.11}$$

with the convention that $\mathcal{M}_{0,t}(H)=0$ if $\mathcal{J}_{0,t}=\emptyset$. In Lemma 3.9, we see that if $\mathbf{m}_H$ is diffuse and supported by the set of leaves of $\mathcal{T}_H$, then there is a measurable way to recover (t,H) from $\mathcal{M}_{0,t}(H)$.

For all $t_1\geq t_0\geq 0$, we also set

$$\mathcal{M}_{t_0,t_1}(H) := \mathcal{M}_{0,t_1-t_0}\big(H^{[t_0]}\big) =: \sum_{a\in\mathcal{J}_{t_0,t_1}} \delta_{(a,\overleftarrow{H}^a,\overrightarrow{H}^a)}\ . \tag{3.12}$$

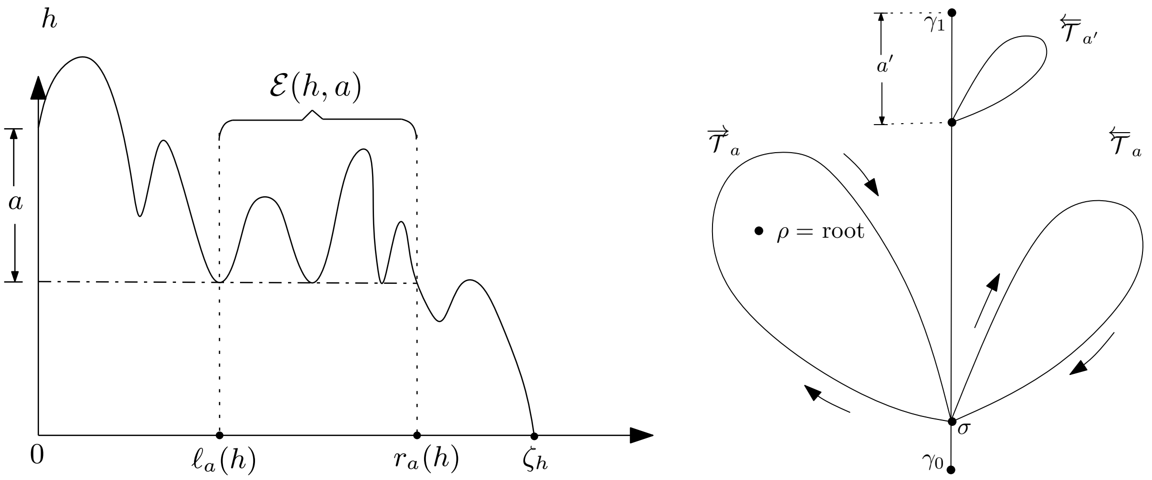

Figure 3.1 – *The figure on the left hand side illustrates the definition of $\mathcal{E}(h,a)$. The figure on the right hand side represents the spinal decomposition of H at times t_0 and t_1 in terms of the tree $\mathcal{T}$ coded by H.*

This point measure on $[0, d_H(t_0,t_1)] \times \mathbb{C}(\mathbb{R}_+, \mathbb{R}_+)^2$ is the *spinal decomposition of H between t_0 and t_1*.

Let us interpret this decomposition in terms of the tree $\mathcal{T}_H$ (see Figure 3.1). Set $\gamma_0 = p_H(t_0)$ and $\gamma_1 = p_H(t_1)$; to simplify, we assume that γ_0 and γ_1 are leaves. Recall that $[\![\gamma_0, \gamma_1]\!]$ is the geodesic path joining γ_0 and γ_1; then $\mathcal{J}_{t_0,t_1} = \{d(\sigma, \gamma_1); \sigma \in \mathtt{Br}(\mathcal{T}_H) \cap [\![\gamma_0, \gamma_1]\!]\}$. For any positive $a \in \mathcal{J}_{t_0,t_1}$, there exists $\sigma \in \mathtt{Br}(\mathcal{T}_H) \cap [\![\gamma_0, \gamma_1]\!]$ such that the following holds true.

- $\overrightarrow{\mathcal{T}}_a := \{\sigma\} \cup \big\{\sigma' \in \mathcal{T}_H : \gamma_0 <_H \sigma' <_H \gamma_1 \text{ and } [\![\gamma_0, \sigma]\!] = [\![\gamma_0, \sigma']\!] \cap [\![\gamma_0, \gamma_1]\!]\big\}$ is the tree grafted at σ on the right hand side of $[\![\gamma_0, \gamma_1]\!]$ and the tree $(\overrightarrow{\mathcal{T}}_a, d, \sigma)$ is coded by $\overrightarrow{H}^a$.
- $\overleftarrow{\mathcal{T}}_a := \{\sigma\} \cup \big\{\sigma' \in \mathcal{T}_H : \text{either } \sigma' <_H \gamma_0 \text{ or } \gamma_1 <_H \sigma' \text{ and } [\![\gamma_0, \sigma]\!] = [\![\gamma_0, \sigma']\!] \cap [\![\gamma_0, \gamma_1]\!]\big\}$ is the tree grafted at σ on the left hand side of $[\![\gamma_0, \gamma_1]\!]$ and the tree $(\overleftarrow{\mathcal{T}}_a, d, \sigma)$ is coded by $\overleftarrow{H}^a$.

Height process and Lévy trees. The Brownian tree (also called Continuum Random Tree) has been introduced by Aldous [8–10]; this model has been extended by Le Gall & Le Jan: in [83], they define the *height process* (further studied by D. & Le Gall [51]) that is the coding function of Lévy trees. Lévy trees appear as scaling limits of Galton-Watson trees and they are the genealogical structure of continuous state branching processes. Let us briefly recall here the definition of the height process and that of Lévy trees.

The law of the height process is characterized by a function $\Psi : \mathbb{R}_+ \to \mathbb{R}_+$ called *branching mechanism*; we shall restrict our attention to the critical and subcritical cases, namely when the branching mechanism Ψ is of the following Lévy-Khintchine form:

$$\forall \lambda \in \mathbb{R}_+, \quad \Psi(\lambda) = \alpha\lambda + \beta\lambda^2 + \int_{(0,\infty)} \big(e^{-\lambda r} - 1 + \lambda r\big)\, \pi(dr)\,, \tag{3.13}$$

where $\alpha, \beta \in \mathbb{R}_+$ and where π is the Lévy measure on $(0,\infty)$ that satisfies $\int_{(0,\infty)} (r \wedge r^2)\, \pi(dr) < \infty$. The height process is derived from a spectrally positive Lévy process whose Laplace exponent is Ψ. It shall be convenient to work with the canonical process $X = (X_t)_{t \geq 0}$ on the space of càdlàg functions $\mathbf{D}(\mathbb{R}_+, \mathbb{R})$ equipped with the Skorohod topology. Let us denote by $\mathbf{P}$ the law of a spectrally Lévy process starting from 0 and whose Laplace exponent is Ψ. Namely,

$$\forall t, \lambda \in \mathbb{R}_+, \quad \mathbb{E}\left[\exp\left(-\lambda X_t\right)\right] = \exp\big(t\Psi(\lambda)\big)\,.$$

Note that the form (3.13) ensures that X under $\mathbf{P}$ does not drift to ∞: see for instance Bertoin [22], Chapter VII for more details. Under the following assumption:

$$\int_1^\infty \frac{d\lambda}{\Psi(\lambda)} < \infty, \tag{3.14}$$

Le Gall & Le Jan [83] (see also D. & Le Gall [51]) have proved that there exists a continuous process $H = (H_t)_{t\geq 0}$ such that for all $t \in \mathbb{R}_+$, the following limit holds in $\mathbf{P}$-probability:

$$H_t = \lim_{\epsilon \to 0} \frac{1}{\epsilon} \int_0^t ds \, \mathbf{1}_{\{I_t^s < X_s < I_t^s + \epsilon\}}, \tag{3.15}$$

where $I_t^s := \inf_{s<r<t} X_r$. The process H is called the Ψ*-height process*. In the Brownian case, namely when $\Psi(\lambda) = \lambda^2$, easy arguments show that H is distributed as a reflected Brownian motion. Le Gall & Le Jan [83] have proved a Ray-Knight theorem for H, which shows that the height process H codes the genealogy of continuous state branching processes (see also D. & Le Gall [51], Theorem 1.4.1). Moreover, the Ψ-height process H appears as the scaling limit of the discrete height process and the contour function of Galton-Watson discrete trees: see D. & Le Gall [51], Chaper 2, for more details.

For all $x \in (0, \infty)$, we set $T_x = \inf\{t \in \mathbb{R}_+ : X_t = -x\}$ that is $\mathbf{P}$-a.s. finite since X under $\mathbf{P}$ does not drift to ∞. We next introduce the following law $\mathbf{P}^x$ on $\mathbb{C}(\mathbb{R}_+, \mathbb{R}_+)$:

$$\mathbf{P}^x \text{ is the law of } (H_{t\wedge T_x})_{t\geq 0} \text{ under } \mathbf{P}, \tag{3.16}$$

The tree $\mathcal{T}_H$ under $\mathbf{P}^x(dH)$ is called the Ψ*-Lévy forest starting from a population of size* x. Then, the mass measure of $\mathcal{T}_H$ under $\mathbf{P}^x(dH)$ satisfies the following important properties:

$$\mathbf{P}^x(dH)\text{-a.s. } \mathbf{m}_H \text{ is diffuse and } \mathbf{m}_H(\mathcal{T}_H \backslash \mathtt{Lf}(\mathcal{T}_H)) = 0, \tag{3.17}$$

where we recall from (3.5) that $\mathtt{Lf}(\mathcal{T}_H)$ stands for the set of leaves of the tree $\mathcal{T}_H$. The Ψ-Lévy forest $(\mathcal{T}_H, d_H, \varrho_H, \mathbf{m}_H)$ is therefore a *continuum tree* according to the definition of Aldous [10].

Each excursion above 0 of H under $\mathbf{P}^x$ corresponds to a tree of the Lévy forest. Let us make this point precise by introducing a Poisson decomposition of H into excursions above 0. To that end, denote by I the infimum process of X:

$$\forall t \in \mathbb{R}_+, \quad I_t = \inf_{0\leq r\leq t} X_r \; .$$

Observe that (3.14) entails that either

$$\beta > 0 \quad \text{or} \quad \int_{(0,1)} r\, \pi(dr) = \infty \; , \tag{3.18}$$

which is equivalent for the Lévy process X to have unbounded variation sample paths; basic results of fluctuation theory (see Bertoin [22], Sections VI.1) entail that $X - I$ is a strong Markov process in $[0, \infty)$ and that 0 is regular for $(0, \infty)$ and recurrent with respect to this Markov process. Moreover, $-I$ is a local time at 0 for $X - I$ (see Bertoin [22], Theorem VII.1). We denote by $\mathbf{N}$ the corresponding excursion measure of $X - I$ above 0.

It is not difficult to derive from (3.15) that H_t only depends on the excursion of $X - I$ above 0 which straddles t. Moreover, we get $\{t \in \mathbb{R}_+ : H_t > 0\} = \{t \in \mathbb{R}_+ : X_t > I_t\}$ and if we denote by (a_i, b_i), $i \in \mathcal{I}$, the connected components of this set and if we set $H_s^i = H_{(a_i+s)\wedge b_i}$, $s \in \mathbb{R}_+$, then the point measure

$$\sum_{i\in\mathcal{I}} \delta_{(-I_{a_i}, H^i)} \tag{3.19}$$

is a Poisson point measure on $\mathbb{R}_+ \times \mathbb{C}(\mathbb{R}_+, \mathbb{R}_+)$ with intensity $dx\, \mathbf{N}(dH)$, where, with a slight abuse of notation, $\mathbf{N}(dH)$ stands for the "distribution" of $H(X)$ under $\mathbf{N}(dX)$. In the Brownian case, up to scaling, $\mathbf{N}$ is Itô positive excursion of Brownian motion and the decomposition (3.19) corresponds to the Poisson decomposition of a reflected Brownian motion above 0.

In what follows, we shall mostly work with the Ψ-height process H under its excursion $\mathbf{N}$ that is a sigma-finite measure on $\mathbb{C}(\mathbb{R}_+, \mathbb{R}_+)$. We simply denote by ζ the *lifetime* of H under $\mathbf{N}$ and we easily check that

$$\mathbf{N}\text{-a.e.} \quad \zeta < \infty\,, \quad H_0 = H_\zeta = 0 \quad \text{and} \quad H_t > 0 \iff t \in (0, \zeta)\,. \tag{3.20}$$

Also note that X and H under $\mathbf{N}$ have the same lifetime ζ and basic results of fluctuation theory (see Bertoin [22], Chapter VII) also entail the following:

$$\forall \lambda \in (0, \infty) \quad \mathbf{N}\big[1 - e^{-\lambda \zeta}\big] = \Psi^{-1}(\lambda), \tag{3.21}$$

where Ψ^{-1} stands for the inverse function of Ψ.

Note that (3.20) shows that H under $\mathbf{N}$ is a coding function as defined above. D. & Le Gall [52] then define the Ψ*-Lévy tree* as the real tree coded by H under $\mathbf{N}$.

Convention. When there there is no risk of confusion, we simply write

$$(\mathcal{T}, d, \rho, \mathbf{m}, \leq, p, \Gamma, D) := (\mathcal{T}_H, d_H, \rho_H, \mathbf{m}_H, \leq_H, p_H, \Gamma_H, D_H)$$

when H is considered under $\mathbf{N}$, $\mathbf{P}^x$ or under other measures on $\mathbb{C}(\mathbb{R}_+, \mathbb{R}_+)$. $\square$

Recall from (3.5) that $\mathtt{Lf}(\mathcal{T})$ stands for the set of leaves of $\mathcal{T}$. Then the mass measure has the following properties:

$$\mathbf{N}\text{-a.e. } \mathbf{m} \text{ is diffuse and } \mathbf{m}(\mathcal{T} \backslash \mathtt{Lf}(\mathcal{T})) = 0. \tag{3.22}$$

Then the Ψ-Lévy tree $(\mathcal{T}, d, \rho, \mathbf{m})$ is therefore a continuum tree according to the definition of Aldous [8].

Diameter decomposition. Recall from (3.4) the definition of the total height Γ and that of the diameter D. Let first briefly recall results on the total height. One checks that the total height is $\mathbf{N}$-a.s. realized at a unique time (see D. & Le Gall [52] and also Abraham & Delmas [3]). Namely,

$$\mathbf{N}\text{-a.e. there exists a unique } \tau \in [0, \zeta] \text{ such that } H_\tau = \Gamma\,. \tag{3.23}$$

Moreover, the distribution of the total height Γ under $\mathbf{N}$ is characterized as follows:

$$\forall t \in (0, \infty), \quad v(t) := \mathbf{N}(\Gamma > t) \quad \text{satisfies} \quad \int_{v(t)}^{\infty} \frac{d\lambda}{\Psi(\lambda)} = t\,. \tag{3.24}$$

Note that $v : (0, \infty) \to (0, \infty)$ is a bijective decreasing C^∞ function and (3.24) implies that on $(0, \infty)$, $\mathbf{N}(\Gamma \in dt) = \Psi(v(t))\, dt$.

Recall from (3.16) that $\mathbf{P}^x$ is the law of $(H_{t \wedge T_x})_{t \geq 0}$ under $\mathbf{P}$, where $T_x = \inf\{t \in \mathbb{R}_+ : X_t = -x\}$. The Poisson decomposition (3.19) implies that $\sup_{t \in [0, T_x]} H_t = \max\{\Gamma(H^i); i \in \mathcal{I} : -I_{a_i} \leq x\}$ and since Γ under $\mathbf{N}$ has a density, then (3.23) and (3.24) entail that

$$\mathbf{P}^x\text{-a.s. there is a unique } \tau \in [0, \zeta] \text{ such that } H_\tau = \Gamma \quad \text{and} \quad \mathbf{P}^x(\Gamma \leq t) = e^{-x v(t)},\ t \in \mathbb{R}_+. \tag{3.25}$$

In [3], Abraham & Delmas generalize Williams' decomposition of the Brownian excursion to the excursion of the Ψ-height process: they first make sense of the conditioned law $\mathbf{N}(\,\cdot\,|\,\Gamma = r)$. Namely they prove that $\mathbf{N}(\,\cdot\,|\,\Gamma = r)$-a.s. $\Gamma = r$, that $r \mapsto \mathbf{N}(\,\cdot\,|\,\Gamma = r)$ is weakly continuous on $\mathbb{C}(\mathbb{R}_+, \mathbb{R}_+)$ and that

$$\mathbf{N} = \int_0^\infty \mathbf{N}(\Gamma \in dr)\, \mathbf{N}(\,\cdot\,|\,\Gamma = r)\,. \tag{3.26}$$

Moreover they provide a Poisson decomposition along the total height of the process: see Section 3.2.2 where a more precise statement is recalled. The first two results of our article provide a similar result for the diameter D of the Ψ-Lévy tree under $\mathbf{N}$. Recall that $p : [0, \zeta] \to \mathcal{T}$ stands for the canonical projection.

Theorem 3.1. *Let Ψ be a branching mechanism of the form (3.13) that satisfies (3.14). Let $\mathcal{T}$ be the Ψ-Lévy tree that is coded by the Ψ-height process H under the excursion measure $\mathbf{N}$ as defined above. Then, the following holds true $\mathbf{N}$-a.e.*

(i) There exists a unique pair $\tau_0, \tau_1 \in [0, \zeta]$ such that $\tau_0 < \tau_1$ and $D = d(\tau_0, \tau_1)$. Moreover, either $H_{\tau_0} = \Gamma$ or $H_{\tau_1} = \Gamma$. Namely, either $\tau_0 = \tau$ or $\tau_1 = \tau$, where τ is the unique time realizing the total height as defined by (3.23).

(ii) Set $\gamma_0 = p(\tau_0)$ and $\gamma_1 = p(\tau_1)$. Then γ_0 and γ_1 are leaves of $\mathcal{T}$. Let γ_{mid} be the mid-point of $[\![\gamma_0, \gamma_1]\!]$: namely, γ_{mid} is the unique point of $[\![\gamma_0, \gamma_1]\!]$ such that $d(\gamma_0, \gamma_{\mathrm{mid}}) = D/2$. Then, there are exactly two times $0 \le \tau^-_{\mathrm{mid}} < \tau^+_{\mathrm{mid}} \le \zeta$ such that $p(\tau^-_{\mathrm{mid}}) = p(\tau^+_{\mathrm{mid}}) = \gamma_{\mathrm{mid}}$, and γ_{mid} is a simple point of $\mathcal{T}$: namely, it is neither a branching point nor a leaf of $\mathcal{T}$.

(iii) For all $r \in (0, \infty)$, we get

$$\mathbf{N}\big(D > 2r\big) = v(r) - \Psi\big(v(r)\big)^2 \int_{v(r)}^{\infty} \frac{d\lambda}{\Psi(\lambda)^2} \, . \tag{3.27}$$

This implies that $\mathbf{N}(D \in dr) = \varphi(r)dr$ on $(0, \infty)$ where the density $\varphi : (0, \infty) \to (0, \infty)$ is given by

$$\forall r \in (0, \infty), \quad \varphi(2r) = \Psi(v(r)) - \Psi(v(r))^2\, \Psi'(v(r)) \int_{v(r)}^{\infty} \frac{d\lambda}{\Psi(\lambda)^2} \, . \tag{3.28}$$

The second main result of our paper is a Poisson decomposition of the subtrees of $\mathcal{T}$ grafted on the diameter $[\![\gamma_0, \gamma_1]\!]$. This result is stated in terms of coding functions and we first need to introduce the following notation: let $H, H' \in \mathbb{C}(\mathbb{R}_+, \mathbb{R}_+)$ be two coding functions as defined above; the *concatenation* of H and H' is the coding function denoted by $H \oplus H'$ and given by

$$\forall t \in \mathbb{R}_+, \qquad (H \oplus H')_t = H_t \quad \text{if } t \in [0, \zeta_H] \quad \text{and} \quad (H \oplus H')_t = H'_{t-\zeta_H} \quad \text{if } t \ge \zeta_H. \tag{3.29}$$

Moreover, to simplify notation we write the following:

$$\forall r \in (0, \infty), \quad \mathbf{N}^{\Gamma}_r = \mathbf{N}(\,\cdot \,|\, \Gamma = r) \, . \tag{3.30}$$

Theorem 3.2. *Let Ψ be a branching mechanism of the form (3.13) that satisfies (3.14). For all $r \in (0, \infty)$, we denote by $\mathbf{Q}_r$ the law on $\mathbb{C}(\mathbb{R}_+, \mathbb{R}_+)$ of $H \oplus H'$ under $\mathbf{N}^{\Gamma}_{r/2}(dH)\mathbf{N}^{\Gamma}_{r/2}(dH')$, where $\mathbf{N}^{\Gamma}_{r/2}$ is defined by (3.30). Namely, for all measurable functions $F : \mathbb{C}(\mathbb{R}_+, \mathbb{R}_+) \to \mathbb{R}_+$,*

$$\mathbf{Q}_r\big[F(H)\big] = \iint_{\mathbb{C}(\mathbb{R}_+, \mathbb{R}_+)^2} \mathbf{N}^{\Gamma}_{r/2}(dH)\mathbf{N}^{\Gamma}_{r/2}(dH')\, F\big(H \oplus H'\big) \, . \tag{3.31}$$

Then $\mathbf{Q}_r$ satisfies the following properties.

(i) $\mathbf{Q}_r$-a.s. $D = r$ and there exists a unique pair of points $\tau_0, \tau_1 \in [0, \zeta]$ such that $D = d(\tau_0, \tau_1)$.

(ii) For all $r \in (0, \infty)$, $\mathbf{Q}_r[\,\zeta\,] = 2\mathbf{N}^{\Gamma}_{r/2}[\,\zeta\,] \in (0, \infty)$. Moreover, the application $r \mapsto \mathbf{Q}_r$ is weakly continuous and for all measurable functions $F : \mathbb{C}(\mathbb{R}_+, \mathbb{R}_+) \to \mathbb{R}_+$ and $f : \mathbb{R}_+ \to \mathbb{R}_+$,

$$\mathbf{N}\big[f(D)F(H)\big] = \int_0^{\infty} \frac{\mathbf{N}(D \in dr)}{\mathbf{Q}_r[\,\zeta\,]}\, f(r)\, \mathbf{Q}_r\Big[\int_0^{\zeta} F\big(H^{[t]}\big)\, dt\Big] \, , \tag{3.32}$$

where $H^{[t]}$ is defined by (3.8).

(iii) Recall the notation τ^-_{mid} and τ^+_{mid} from Theorem 3.1 (ii). Then, for all $r\in(0,\infty)$,

$$\mathbf{N}\big[F\big(H^{[\tau^-_{\mathrm{mid}}]}\big)\,\big|\,D=r\big] = \frac{1}{\mathbf{N}^\Gamma_{r/2}[\,\zeta\,]}\iint_{\mathbb{C}(\mathbb{R}_+,\mathbb{R}_+)^2}\mathbf{N}^\Gamma_{r/2}(dH)\mathbf{N}^\Gamma_{r/2}(dH')\;\zeta_{H'}F\big(H\oplus H'\big)\,, \tag{3.33}$$

where $\mathbf{N}(\,\cdot\,\big|\,D=r)$ makes sense for all $r\in(0,\infty)$ thanks to (3.32).

(iv) Recall from (3.16) the notation $\mathbf{P}^y$. To simplify notation, we write for all $y,b\in(0,\infty)$

$$\mathbf{N}_b=\mathbf{N}\big(\,\cdot\cap\{\Gamma\le b\}\big)\quad\textit{and}\quad \mathbf{P}^y_b=\mathbf{P}^y\big(\,\cdot\cap\{\Gamma\le b\}\big), \tag{3.34}$$

Then, under $\mathbf{Q}_r$, $\mathcal{M}_{\tau_0,\tau_1}(da\,d\overleftarrow{H}\,d\overrightarrow{H})$, defined by (3.12), is a Poisson point measure on $[0,r]\times\mathbb{C}(\mathbb{R}_+,\mathbb{R}_+)^2$ whose intensity is

$$\begin{aligned}\beta\mathbf{1}_{[0,r]}(a)da\,\Big(\delta_{\mathbf{0}}(d\overleftarrow{H})\mathbf{N}_{a\wedge(r-a)}(d\overrightarrow{H})+\mathbf{N}_{a\wedge(r-a)}(d\overleftarrow{H})\delta_{\mathbf{0}}(d\overrightarrow{H})\Big)&\\ +\;\mathbf{1}_{[0,r]}(a)da\int_{(0,\infty)}\pi(dz)\int_0^z dx\;\mathbf{P}^x_{a\wedge(r-a)}\big(d\overleftarrow{H}\big)\,\mathbf{P}^{z-x}_{a\wedge(r-a)}\big(d\overrightarrow{H}\big),&\end{aligned} \tag{3.35}$$

where β and π are defined in (3.13).

Comment 1. *As already mentioned, the previous theorem makes sense of $\mathbf{N}\big(\,\cdot\,\big|D=r\big)$ and for all measurable functions $F:\mathbb{C}(\mathbb{R}_+,\mathbb{R}_+)\to\mathbb{R}_+$, we have*

$$\forall r\in(0,\infty),\qquad \mathbf{N}\big[\,F(H)\,\big|\,D=r\big]:=\mathbf{Q}_r\Big[\int_0^\zeta F\big(H^{[t]}\big)\,dt\Big]\Big/\mathbf{Q}_r[\,\zeta\,]\,, \tag{3.36}$$

Namely, Theorem 3.2 (i) entails that $\mathbf{N}(\,\cdot\,\big|\,D=r)$-a.s. $D=r$. Then (3.31) combined with the already mentioned continuity of $r\mapsto\mathbf{N}(\,\cdot\,|\,\Gamma=r/2)$ easily implies that $r\mapsto\mathbf{N}(\,\cdot\,\big|\,D=r)$ is weakly continuous on $\mathbb{C}(\mathbb{R}_+,\mathbb{R}_+)$. Moreover, (3.32) can be rewritten:

$$\mathbf{N}=\int_0^\infty\mathbf{N}(D\in dr)\,\mathbf{N}(\,\cdot\,|\,D=r) \tag{3.37}$$

that is the exact analogous of (3.26). We mention that the proof of Theorem 3.2 relies on the decomposition (3.26) due to Abraham & Delmas [3]. □

Comment 2. *It is easy to check from (3.8) that for all t_0,t, $(H^{[t]})^{[t_0]}=H^{[t+t_0]}$. Therefore, (3.32) implies that H under $\mathbf{N}$ is invariant under rerooting. Namely, for all measurable functions $F:\mathbb{C}(\mathbb{R}_+,\mathbb{R}_+)\to\mathbb{R}_+$,*

$$\forall t_0\in\mathbb{R}_+,\quad \mathbf{N}\big[\mathbf{1}_{\{\zeta\ge t_0\}}F\big(H^{[t_0]}\big)\big]=\mathbf{N}\big[\mathbf{1}_{\{\zeta\ge t_0\}}F\big(H\big)\big]\,, \tag{3.38}$$

which is quite close to Proposition 2.1 in D. & Le Gall [53], that is used in the proof of Theorem 3.2. □

Comment 3. *As shown by (3.36), $\mathbf{N}\big(\,\cdot\,\big|D=r\big)$ is derived from $\mathbf{Q}_r$ by a uniform rerooting. This property suggests that the law of the compact real tree $(\mathcal{T},d)$ coded by H under $\mathbf{Q}_r$, without its root, is the scaling limit of natural models of labeled unrooted trees conditioned by their diameter.* □

Comment 4. *Another reason for introducing the law $\mathbf{Q}_r$ is the following: we deduce from (3.36) that for all measurable functions $F:\mathbb{C}(\mathbb{R}_+,\mathbb{R}_+)\to\mathbb{R}_+$,*

$$\mathbf{N}\big[F(H^{[\tau_0]})\,\big|\,D=r\big]=\mathbf{Q}_r\big[\zeta F(H^{[\tau_0]})\big]\big/\mathbf{Q}_r[\,\zeta\,]\,, \tag{3.39}$$

where τ_0 is as in Theorem 3.1. As shown by Theorem 3.2 (iv), H under $\mathbf{Q}_r$ enjoys a Poisson decomposition along its diameter. However (3.39) also implies that this not the case of H under $\mathbf{N}(\,\cdot\,|\,D=r)$.
□

The law of Γ and of D of stable Lévy trees conditioned by their total mass. In application of Theorem 3.2, we compute the law of Γ and D under $\mathbf{N}(\,\cdot\,|\,\zeta = 1)$ in the cases where Ψ is a stable branching mechanism. Namely, we fix $\gamma \in (1, 2]$ and

$$\Psi(\lambda) = \lambda^\gamma, \quad \lambda \in \mathbb{R}_+ \,,$$

that is called the *γ-stable branching mechanism*. We first recall the definition of the law $\mathbf{N}(\,\cdot\,|\,\zeta = 1)$ for such a branching mechanism.

When Ψ is γ-stable, the Lévy process X under $\mathbf{P}$ satisfies the following scaling property: for all $r \in (0, \infty)$, $(r^{-\frac{1}{\gamma}} X_{rt})_{t\geq 0}$ has the same law as X, which easily entails by (3.15) that under $\mathbf{P}$, $(r^{-\frac{\gamma-1}{\gamma}} H_{rt})_{t\geq 0}$ has the same law as H and the Poisson decomposition (3.19) implies the following:

$$\big(r^{-\frac{\gamma-1}{\gamma}} H_{rt}\big)_{t\geq 0} \quad \text{under} \quad r^{\frac{1}{\gamma}}\, \mathbf{N} \quad \overset{\text{(law)}}{=} \quad H \quad \text{under} \quad \mathbf{N} \,. \tag{3.40}$$

We then easily derive from (3.21) that

$$\mathbf{N}(\zeta \in dr) = p_\gamma(r)\, dr \,, \quad \text{where} \quad p_\gamma(r) = c_\gamma r^{-1-\frac{1}{\gamma}} \quad \text{with} \quad 1/c_\gamma = \gamma \Gamma_e\big(\tfrac{\gamma-1}{\gamma}\big) \,. \tag{3.41}$$

Here Γ_e stands for Euler's Gamma function. By (3.40), there exists a family of laws on $\mathbb{C}(\mathbb{R}_+, \mathbb{R}_+)$ denoted by $\mathbf{N}(\,\cdot\,|\,\zeta = r)$, $r \in (0, \infty)$, such that $r \mapsto \mathbf{N}(\,\cdot\,|\,\zeta = r)$ is weakly continous on $\mathbb{C}(\mathbb{R}_+, \mathbb{R}_+)$, such that $\mathbf{N}(\,\cdot\,|\,\zeta = r)$-a.s. $\zeta = r$ and such that

$$\mathbf{N} = \int_0^\infty \mathbf{N}(\,\cdot\,|\,\zeta = r)\, \mathbf{N}(\zeta \in dr) \,. \tag{3.42}$$

Moreover, by (3.40), $\big(r^{-\frac{\gamma-1}{\gamma}} H_{rt}\big)_{t\geq 0}$ under $\mathbf{N}(\,\cdot\,|\,\zeta = r)$ has the same law as H under $\mathbf{N}(\,\cdot\,|\,\zeta = 1)$. We call $\mathbf{N}(\,\cdot\,|\,\zeta = 1)$ the *normalized law of the γ-stable height process* and to simplify notation we set

$$\mathbf{N}_{\text{nr}} := \mathbf{N}(\,\cdot\,|\,\zeta = 1) \tag{3.43}$$

Thus, for all measurable functions $F : \mathbb{C}(\mathbb{R}_+, \mathbb{R}_+) \to \mathbb{R}_+$,

$$\mathbf{N}\big[F(H)\big] = c_\gamma \int_0^\infty dr\, r^{-1-\frac{1}{\gamma}}\, \mathbf{N}_{\text{nr}}\Big[F\Big(\big(r^{\frac{\gamma-1}{\gamma}} H_{t/r}\big)_{t\geq 0}\Big)\Big] \,. \tag{3.44}$$

When $\gamma = 2$, $\mathbf{N}_{\text{nr}}$ is, up to scaling, the normalized Brownian excursion that is, as shown by Aldous [10], the scaling limit of the contour process of the uniform (ordered rooted) tree with n vertices as $n \to \infty$; Aldous [10] also extends this limit theorem to Galton-Watson trees conditioned to have n vertices and whose offspring distribution has a second moment. This result has been extended by D. [49] to Galton-Watson trees conditioned to have n vertices and whose offspring distribution is in the domain of attraction of a γ-stable law, the limiting process being in this case the normalized excursion of the γ-stable height process. See also Kortchemski [76] for scaling limits of Galton-Watson tree conditioned to have n leaves.

We next introduce $w : (0, \infty) \to (1, \infty)$ that is the unique C^∞ decreasing bijection that satisfies the following integral equation:

$$\forall y \in (0, \infty), \qquad \int_{w(y)}^\infty \frac{du}{u^\gamma - 1} = y \,. \tag{3.45}$$

We refer to Section 3.3.1 for a probabilistic interpretation of w and further properties. The following proposition characterizes the joint law of Γ and D under $\mathbf{N}_{\text{nr}}$ by the mean of Laplace transforms.

Proposition 3.3. *Fix $\gamma \in (1,2]$ and $\Psi(\lambda) = \lambda^\gamma$, $\lambda \in \mathbb{R}_+$. Recall from (3.43) the definition of the law $\mathbf{N}_{\mathrm{nr}}$ of the normalized excursion of the γ-stable height process. We then set*

$$\forall \lambda, y, z \in (0,\infty), \quad \mathrm{L}_\lambda(y,z) := c_\gamma \int_0^\infty e^{-\lambda r} r^{-1-\frac{1}{\gamma}}\, \mathbf{N}_{\mathrm{nr}}\big(r^{\frac{\gamma-1}{\gamma}} D > 2y \,;\, r^{\frac{\gamma-1}{\gamma}} \Gamma > z \big)\, dr \ , \tag{3.46}$$

where we recall from (3.41) that $1/c_\gamma = \gamma \Gamma_e\big(\frac{\gamma-1}{\gamma}\big)$, Γ_e standing for Euler's Gamma function. Note that

$$\forall \lambda, y, z \in (0,\infty), \quad \mathrm{L}_1(y,z) = \lambda^{-\frac{1}{\gamma}} \mathrm{L}_\lambda\big(\lambda^{-\frac{\gamma-1}{\gamma}} y \,,\, \lambda^{-\frac{\gamma-1}{\gamma}} z\big) \ . \tag{3.47}$$

Recall from (3.45) the definition of w. Then,

$$\mathrm{L}_1(y,z) = w(y \vee z) - 1 - \tfrac{1}{\gamma} \mathbf{1}_{\{z<2y\}} \left(w(y)^\gamma - 1 \right)^2 \left(\frac{w\big(y \wedge (2y-z)\big)}{w\big(y \wedge (2y-z)\big)^\gamma - 1} - (\gamma-1)\big(y \wedge (2y-z)\big) \right) . \tag{3.48}$$

In particular, for all $y, z \in (0,\infty)$,

$$\mathrm{L}_1(0,z) = w(z) - 1 \quad \textit{and} \quad \mathrm{L}_1(y,0) = w(y) - 1 - \tfrac{1}{\gamma}\big(w(y)^\gamma - 1\big)\Big(w(y) - (\gamma-1) y \big(w(y)^\gamma - 1\big)\Big). \tag{3.49}$$

The previous proposition is known in the Brownian case, where $w(y) = \coth(y)$: see W. [100]. In the Brownian case, standard computations derived from (3.49) imply the following power expansions that hold true for all $y \in (0,\infty)$:

$$\mathbf{N}_{\mathrm{nr}}\big(\Gamma > y\big) = 2 \sum_{n \geq 1} \big(2n^2 y^2 - 1\big)\, e^{-n^2 y^2} \ , \tag{3.50}$$

$$\mathbf{N}_{\mathrm{nr}}\big(D > y\big) = \sum_{n \geq 2} (n^2 - 1)\big(\tfrac{1}{6} n^4 y^4 - 2n^2 y^2 + 2\big) e^{-n^2 y^2/4} \ . \tag{3.51}$$

See W. [100] for more details.

We next provide similar asymptotic expansions in the non-Brownian stable cases. To that end, we introduce $s_\gamma : (0,\infty) \to (0,\infty)$ as the continuous version of the density of the spectrally positive $\frac{\gamma-1}{\gamma}$-stable distribution; more precisely, s_γ is characterized by the following:

$$\forall \lambda \in \mathbb{R}_+, \qquad \int_0^\infty e^{-\lambda x} s_\gamma(x)\, dx = \exp(-\gamma \lambda^{\frac{\gamma-1}{\gamma}}) \ . \tag{3.52}$$

The following asymptotic expansion of s_γ at 0 is due to Zolotarev (see Theorem 2.5.2 [103]): for all integer $N \geq 1$,

$$\big(2\pi(1 - \tfrac{1}{\gamma})\big)^{\frac{1}{2}} x^{\frac{\gamma+1}{2}} e^{1/x^{\gamma-1}} s_\gamma\big((\gamma-1)x\big) = 1 + \sum_{1 \leq n < N} S_n\, x^{n(\gamma-1)} + \mathcal{O}_{N,\gamma}\big(x^{N(\gamma-1)}\big), \quad \text{as } x \to 0. \tag{3.53}$$

Here $\mathcal{O}_{N,\gamma}$ means that the expansion depends on N and γ. Next, note that S_n depends on n and γ but we skip the dependence in γ to simplify notation.

Remark 21. *In the Brownian case where $\gamma = 2$, it is well-known that*

$$s_2(x) = \pi^{-\frac{1}{2}} x^{-\frac{3}{2}} e^{-1/x}, \quad x \in \mathbb{R}_+$$

Then, $S_0 = 1$ and $S_n = 0$, for all $n \geq 1$. □

For generic $\gamma \in (1,2)$, this asymptotic expansion does not yield a converging power expansion (although it is the case if $\gamma = 2$). See Section 3.4.1 for more details on s_γ. To state our result we first need to introduce an auxiliary function derived from s_γ as follows.

Proposition 3.4. *Let $\gamma \in (1,2]$. Recall from (3.52) the definition of s_γ. We introduce the following function:*

$$\forall x \in \mathbb{R}_+, \quad \theta(x) := (\gamma-1)\, x^{-1} s_\gamma(x) - \tfrac{\gamma-1}{\gamma} x^{-1-\frac{1}{\gamma}} \int_0^x dy\, y^{\frac{1}{\gamma}-1} s_\gamma(y)\ . \tag{3.54}$$

Then, the following holds true.

(i) θ is well-defined, continuous,

$$\int_0^\infty dx\, |\theta(x)| < \infty \quad \textit{and} \quad \int_0^\infty dx\, e^{-\lambda x}\theta(x) = \lambda^{\frac{1}{\gamma}} e^{-\gamma \lambda^{\frac{\gamma-1}{\gamma}}}, \quad \lambda \in \mathbb{R}_+. \tag{3.55}$$

(ii) Recall from (3.53) the definition of the sequence $(S_n)_{n\geq 0}$, with $S_0=1$. Let $(V_n)_{n\geq 0}$ be a sequence of real numbers recursively defined by $V_0=1$ and

$$\forall n \in \mathbb{N}, \qquad V_{n+1} = S_{n+1} + \big(n - \tfrac{1}{2} - \tfrac{1}{\gamma-1}\big) S_n - \big(n - \tfrac{1}{2} - \tfrac{1}{\gamma}\big) V_n\ . \tag{3.56}$$

Then, for all integer $N \geq 1$,

$$\big(2\pi(1-\tfrac{1}{\gamma})\big)^{\frac{1}{2}}\, x^{\frac{\gamma+3}{2}} e^{1/x^{\gamma-1}} \theta\big((\gamma-1)x\big) = 1 + \sum_{1\leq n<N} V_n\, x^{n(\gamma-1)} \ + \mathcal{O}_{N,\gamma}\big(x^{N(\gamma-1)}\big), \tag{3.57}$$

as $x \to 0$.

We use θ to get the asymptotic expansion of the law of the total height of the normalized γ-stable tree as follows.

Theorem 3.5. *Let $\gamma \in (1,2]$. We introduce the following function:*

$$\forall r \in \mathbb{R}_+, \quad \xi(r) := r^{-\frac{\gamma+1}{\gamma-1}} \theta\big(r^{-\frac{\gamma}{\gamma-1}}\big)\ . \tag{3.58}$$

where θ is defined in (3.54). Then, there exists a real valued sequence $(\beta_n)_{n\geq 1}$ and $x_1 \in (0,1)$ such that

$$\sum_{n\geq 1} |\beta_n| x_1^n \ < \infty \quad \textit{and} \quad \forall r \in (0,\infty), \quad \sum_{n\geq 1} |\beta_n| \sup_{s\in[r,\infty)} |\xi(ns)| \ < \infty\ , \tag{3.59}$$

and such that

$$\forall r \in (0,\infty), \quad c_\gamma \mathbf{N}_{\mathrm{nr}}\big(\Gamma > r\big) = \sum_{n\geq 1} \beta_n\, \xi(nr)\ , \tag{3.60}$$

where we recall from (3.41) that $1/c_\gamma = \gamma \Gamma_e\big(\frac{\gamma-1}{\gamma}\big)$, Γ_e standing for Euler's gamma function. Moreover, for all integers $N \geq 1$, as $r \to \infty$,

$$\frac{1}{C_1} r^{-1-\frac{\gamma}{2}} e^{r^\gamma}\, \mathbf{N}_{\mathrm{nr}}\Big(\Gamma > r(\gamma-1)^{-\frac{\gamma-1}{\gamma}}\Big) \ = \ 1 + \sum_{1\leq n<N} V_n\, r^{-n\gamma} \ + \mathcal{O}_{N,\gamma}\big(r^{-N\gamma}\big)\ , \tag{3.61}$$

where $C_1 := (2\pi)^{-\frac{1}{2}}(\gamma-1)^{\frac{1}{2}+\frac{1}{\gamma}} \gamma^{\frac{3}{2}}\, \Gamma_e(\frac{\gamma-1}{\gamma}) \exp(C_0)$, where

$$C_0 := \gamma \int_1^\infty \frac{du}{(u+1)^\gamma - 1} - \int_0^1 \frac{du}{u} \, \frac{(u+1)^\gamma - 1 - \gamma u}{(u+1)^\gamma - 1}\ , \tag{3.62}$$

and where the sequence $(V_n)_{n\geq 1}$ is recursively defined by (3.56) in Proposition 3.4.

Remark 22. *The convergence in (3.60) is rapid. Indeed, by (3.57), we see that $\xi(nr)$ is of order*

$$(nr)^{1+\frac{\gamma}{2}} \exp(-n^\gamma(\gamma-1)^{\frac{1}{\gamma-1}} r^\gamma)\ .$$

Then, the asymptotic expansion (3.61) is that of the first term of (3.60) that is $c_\gamma^{-1}\beta_1\, \xi(r)$. □

Remark 23. *The definition of the sequence* $(\beta_n)_{n\geq 0}$ *is involved: see Lemma 3.24 and its proof for a precise definition. However, in the Brownian case, everything can be explicitly computed: for all* $n\geq 1$, $\beta_n=2$, $\xi(r)=(4\pi)^{-\frac{1}{2}}(2r^2-1)e^{-r^2}$, $c_2=(4\pi)^{-\frac{1}{2}}$, *and we recover (3.50) from (3.60); moreover,* $C_0=\log 2$, $C_1=4$, $V_0=1$, $V_1=-\frac{1}{2}$ *and* $V_n=0$, *for all* $n\geq 2$. □

To state the result concerning the diameter, we need precise results on the derivative of the $\frac{\gamma-1}{\gamma}$-stable density.

Proposition 3.6. *Let* $\gamma\in(1,2]$. *Recall from (3.52) the definition of the density* s_γ. *Then* s_γ *is* C^1 *on* $\mathbb{R}_+$,

$$\int_0^\infty dx\,|s'_\gamma(x)|<\infty \qquad \text{and} \qquad \int_0^\infty dx\,e^{-\lambda x}s'_\gamma(x)=\lambda e^{-\gamma\lambda^{\frac{\gamma-1}{\gamma}}},\quad \lambda\in\mathbb{R}_+. \tag{3.63}$$

Moreover, s'_γ *has the following asymptotic expansion: recall from (3.53) the definition of the sequence* $(S_n)_{n\geq 0}$, *with* $S_0=1$; *let* $(T_n)_{n\geq 0}$ *be a sequence of real numbers recursively defined by* $T_0=1$ *and*

$$\forall n\in\mathbb{N},\qquad T_{n+1}:=S_{n+1}+\big(n-\tfrac{1}{2}-\tfrac{1}{\gamma-1}\big)S_n\,. \tag{3.64}$$

Then, for all positive integers N, *we have*

$$\big(2\pi(1-\tfrac{1}{\gamma})\big)^{\frac{1}{2}}\,x^{\frac{3\gamma+1}{2}}e^{1/x^{\gamma-1}}s'_\gamma\big((\gamma-1)x\big)=1+\sum_{1\leq n<N}T_n\,x^{n(\gamma-1)}\;+\mathcal{O}_{N,\gamma}\big(x^{N(\gamma-1)}\big), \tag{3.65}$$

as $x\to 0$.

The asymptotic expansion of the law of the diameter of the normalized γ-stable tree is then given in the following theorem.

Theorem 3.7. *Let* $\gamma\in(1,2]$. *Recall from (3.58) the definition of the function* ξ. *We also introduce the following function:*

$$\forall r\in\mathbb{R}_+,\quad \overline{\xi}(r):=r^{-\frac{\gamma+1}{\gamma-1}}s'_\gamma\big(r^{-\frac{\gamma}{\gamma-1}}\big)\,, \tag{3.66}$$

where s'_γ *is the derivative of the density* s_γ *defined in (3.52). Then there exist two real valued sequences* $(\gamma_n)_{n\geq 2}$ *and* $(\delta_n)_{n\geq 2}$ *and* $x_2\in(0,1)$ *such that*

$$\sum_{n\geq 2}(|\gamma_n|+|\delta_n|)x_2^n<\infty \quad \text{and} \quad \forall r\in(0,\infty),\quad \sum_{n\geq 2}|\gamma_n|\sup_{s\in[r,\infty)}|\overline{\xi}(ns)|+|\delta_n|\sup_{s\in[r,\infty)}|\xi(ns)|<\infty\,, \tag{3.67}$$

and such that

$$\forall r\in(0,\infty),\quad c_\gamma\mathbf{N}_{\mathrm{nr}}\big(D>2r\big)=\sum_{n\geq 2}\gamma_n\overline{\xi}(nr)+\delta_n\xi(nr)\,, \tag{3.68}$$

where we recall from (3.41) that $1/c_\gamma=\gamma\Gamma_e\big(\frac{\gamma-1}{\gamma}\big)$, Γ_e *standing for Euler's gamma function. Moreover, for all integers* $N\geq 1$, *as* $r\to\infty$,

$$\frac{1}{C_2}\,r^{-1-\frac{3\gamma}{2}}\,e^{r^\gamma}\,\mathbf{N}_{\mathrm{nr}}\Big(D>r(\gamma-1)^{-\frac{\gamma-1}{\gamma}}\Big)\;=\;1+\sum_{1\leq n<N}U_n\,r^{-n\gamma}\;+\mathcal{O}_{\gamma,N}\big(r^{-N\gamma}\big)\,, \tag{3.69}$$

where $C_2:=(8\pi)^{-\frac{1}{2}}(\gamma-1)^{\frac{3}{2}+\frac{1}{\gamma}}\gamma^{\frac{5}{2}}\,\Gamma_e\big(\frac{\gamma-1}{\gamma}\big)\exp(2C_0)$, *where* C_0 *is defined by (3.62) and where the sequence* $(U_n)_{n\geq 1}$ *is recursively defined by* $U_0=1$ *and*

$$\forall n\geq 1\,,\qquad U_n=T_n-\tfrac{\gamma+1}{\gamma(\gamma-1)}V_{n-1}\,. \tag{3.70}$$

Here $(T_n)_{n\geq 0}$ *is defined by (3.64) and* $(V_n)_{n\geq 0}$ *is defined by (3.56).*

Remark 24. *The convergence in (3.68) is rapid. Indeed, by (3.65) and (3.57) we see that $\overline{\xi}(nr/2)$ and $\xi(nr/2)$ are of respective order*

$$(nr)^{1+\frac{3\gamma}{2}}\exp(-n^\gamma 2^{-\gamma}(\gamma-1)^{\frac{1}{\gamma-1}}r^\gamma) \quad \textit{and} \quad (nr)^{1+\frac{\gamma}{2}}\exp(-n^\gamma 2^{-\gamma}(\gamma-1)^{\frac{1}{\gamma-1}}r^\gamma)\ .$$

Then the asymptotic expansion (3.69) is that of $c_\gamma^{-1}\gamma_2\,\overline{\xi}(r)+c_\gamma^{-1}\delta_2\,\xi(r)$. $\square$

Remark 25. *The definitions of the sequences $(\gamma_n)_{n\geq 0}$ and $(\delta_n)_{n\geq 0}$ are involved: see the proof of Lemma 3.25 for a precise definition. However, in the Brownian case, everything can be computed explicitly:*

$$\forall n\geq 2,\quad \gamma_n=\tfrac{4}{3}(n^2-1),\quad \delta_n=-2(n^2-1)\quad \textit{and}\quad \overline{\xi}(r)=\pi^{-\frac{1}{2}}r^2\big(r^2-\tfrac{3}{2}\big)e^{-r^2}\ ,$$

which allows to recover (3.51) from (3.68). Moreover, $C_2=8$, $U_0=1$, $U_1=-3$, $U_2=-\frac{3}{4}$ and $U_n=0$, for all $n\geq 3$. $\square$

The paper is organized as follows. Section 3.2 is devoted to the proof of Theorem 3.1 and of Theorem 3.2: in Section 3.2.1, we discuss an important geometric property of the diameter of real trees (Lemma 3.8) and we explain the spinal decomposition according to the total height, the result of Abraham & Delmas [3] being recalled in Section 3.2.2 where the proofs of Theorem 3.1 and Theorem 3.2 are actually given. Proposition 3.3, that characterizes the joint law of the total height and the diameter of normalized stable trees, is proved in Section 3.3. Theorem 3.5 and Theorem 3.7 are proved in Section 3.4.

3.2 Proof of the diameter decomposition.

3.2.1 Geometric properties of the diameter of real trees; height decomposition.

In this section we gather deterministic results on real trees and their coding functions: we first prove a key lemma on the diameter of real trees; we next discuss how to reconstruct the coding function H from a spinal decomposition $\mathcal{M}_{0,t}(H)$, under a specific assumption on the mass measure $\mathbf{m}_H$ on $\mathcal{T}_H$; then we discuss a decomposition related to the total height.

Total height and diameter of compact rooted real trees. The following result connects the total height and the diameter of a compact rooted real tree.

Lemma 3.8. *Let (T,d,ρ) be a compact rooted real tree. We denote by Γ and D resp. its total height and its diameter: $\Gamma:=\sup_{\sigma\in T}d(\rho,\sigma)$ and $D=\sup_{\sigma,\sigma'\in T}d(\sigma,\sigma')$. Then, the following holds true.*

(i) There exist $\sigma,\sigma_0,\sigma_1\in T$, such that $\Gamma=d(\rho,\sigma)$ and $D=d(\sigma_0,\sigma_1)$. This entails

$$\Gamma\leq D\leq 2\Gamma\ . \tag{3.71}$$

(ii) Let $\sigma_0,\sigma_1\in T$ be such that $D=d(\sigma_0,\sigma_1)$. Then either $d(\rho,\sigma_0)=\Gamma$ or $d(\rho,\sigma_1)=\Gamma$. More precisely,

$$d(\rho,\sigma_0)\geq d(\rho,\sigma_1)\implies d(\rho,\sigma_0)=\Gamma \qquad \textit{and} \qquad d(\rho,\sigma_1)\geq d(\rho,\sigma_0)\implies d(\rho,\sigma_1)=\Gamma\ . \tag{3.72}$$

Proof. First note that $\gamma\in T\mapsto d(\rho,\gamma)$ and $(\gamma,\gamma')\in T^2\mapsto d(\gamma,\gamma')$ are real valued continuous functions defined on compact spaces; basic topological arguments entail the existence of $\sigma,\sigma_0,\sigma_1\in T$ as in (i). The inequality $\Gamma\leq D$ is an immediate consequence of the definitions of Γ and D. The triangle inequality next entails that $D\leq d(\sigma_0,\rho)+d(\rho,\sigma_1)\leq 2\Gamma$, which completes the proof of (3.71) and of (i).

Let $\sigma, \sigma_0, \sigma_1 \in T$ be as in (i). By the four points inequality (3.1) and basic inequalities, we get

$$\begin{aligned}\Gamma + D = d(\rho,\sigma) + d(\sigma_0,\sigma_1) &\leq \max\big(d(\rho,\sigma_0) + d(\sigma,\sigma_1)\,;\, d(\rho,\sigma_1) + d(\sigma,\sigma_0)\big)\\ &\leq \max\big(d(\rho,\sigma_0); d(\rho,\sigma_1)\big) + \max\big(d(\sigma,\sigma_1); d(\sigma,\sigma_0)\big)\;.\end{aligned}$$

If $\max\big(d(\rho,\sigma_0); d(\rho,\sigma_1)\big) < \Gamma$, then the previous inequality implies that $D < \max\big(d(\sigma,\sigma_1); d(\sigma,\sigma_0)\big)$, which is absurd. Thus, $\max\big(d(\rho,\sigma_0); d(\rho,\sigma_1)\big) = \Gamma$, which easily entails the desired result. ■

Coding functions and their spinal decompositions. Recall that $\mathbf{0}$ stands for the null function of $\mathbb{C}(\mathbb{R}_+,\mathbb{R}_+)$. We denote by $\mathbb{C}_c(\mathbb{R}_+,\mathbb{R}_+)$ the functions of $\mathbb{C}(\mathbb{R}_+,\mathbb{R}_+)$ with compact support.

Definition. We introduce the set of coding functions:

$$\mathtt{Exc} = \big\{ H \in \mathbb{C}_c(\mathbb{R}_+,\mathbb{R}_+) : \; H_0 = 0, \; H \neq \mathbf{0}, \, \mathbf{m}_H \text{ is diffuse and } \; \mathbf{m}_H(\mathcal{T}_H \backslash \mathtt{Lf}(\mathcal{T}_H)) = 0 \big\}\;, \tag{3.73}$$

where we recall from (3.5) that $\mathtt{Lf}(\mathcal{T}_H)$ stands for the set of leaves of $\mathcal{T}_H$ and where we recall from (3.6) that $\mathbf{m}_H$ stands for the mass measure of $\mathcal{T}_H$. Then, we set

$$\mathcal{H} = \big\{ B \cap \mathtt{Exc}\,;\, B \text{ Borel subset of } \mathbb{C}(\mathbb{R}_+,\mathbb{R}_+) \big\}\;. \tag{3.74}$$

that is the trace sigma field on $\mathtt{Exc}$ of the Borel sigma field of $\mathbb{C}(\mathbb{R}_+,\mathbb{R}_+)$. □

Remark 26. *Let $H \in \mathtt{Exc}$ and let $s_0, s_1 \in (0,\zeta_H)$ be such that $s_0 < s_1$ and $d_H(s_0,s_1) = 0$. then, we easily check that $H^{[s_0]}_{\cdot\wedge(s_1-s_0)} \in \mathtt{Exc}$.* □

Remark 27. *Recall from (3.17) and from (3.22) that $\mathbf{P}^x$ and $\mathbf{N}$ are supported by $\mathtt{Exc}$.* □

Definition. We introduce the following subset of $\mathbb{R}_+ \times \mathbb{C}(\mathbb{R}_+,\mathbb{R}_+)^2$:

$$E := \mathbb{R}_+ \times \big(\mathtt{Exc} \times (\mathtt{Exc} \cup \{\mathbf{0}\}) \cup (\mathtt{Exc} \cup \{\mathbf{0}\}) \times \mathtt{Exc}\big) \tag{3.75}$$

and we denote by $\mathscr{M}_{\mathrm{pt}}(E)$ the set of point measures

$$M(da\, d\overleftarrow{H}\, d\overrightarrow{H}) = \sum_{a \in \mathcal{J}} \delta_{(a, \overleftarrow{H}^a, \overrightarrow{H}^a)} \quad \text{on } E$$

that satisfy the following conditions:

$$\begin{aligned}&\exists\, r \in \mathbb{R}_+ \text{ such that the closure of } \mathcal{J} \text{ is } [0,r] \quad \text{and} \quad \forall \varepsilon, \eta \in (0,\infty),\\ &\qquad\qquad \#\big\{ a \in \mathcal{J} : \; \Gamma(\overleftarrow{H}^a) \vee \Gamma(\overrightarrow{H}^a) > \eta \quad \text{or} \quad \zeta_{\overleftarrow{H}^a} \vee \zeta_{\overrightarrow{H}^a} > \varepsilon \big\} < \infty\;.\end{aligned} \tag{3.76}$$

We then equip $\mathscr{M}_{\mathrm{pt}}(E)$ with the sigma field $\mathcal{G}$ generated by the applications $M \in \mathscr{M}_{\mathrm{pt}}(E) \mapsto M(A)$, where A ranges among the Borel subsets of $\mathbb{R}_+ \times \mathbb{C}(\mathbb{R}_+,\mathbb{R}_+)^2$. □

The following lemma, whose proof is postponed in Appendix, asserts that H can be recovered in a measurable way from the spinal decomposition $\mathcal{M}_{0,t}(H)$, as defined in (3.11).

Lemma 3.9. *Recall from above the definition of the measurable spaces $(\mathtt{Exc},\mathcal{H})$ and $(\mathscr{M}_{\mathrm{pt}}(E),\mathcal{G})$. The the following holds true.*

(i) For all $t \in (0,\infty)$, we set $\{\zeta > t\} := \{H \in \mathtt{Exc} : \zeta_H > t\}$. Then, $\{\zeta > t\} \in \mathcal{H}$ and

$$H \in \{\zeta > t\} \longmapsto \mathcal{M}_{0,t}(H) \in \mathscr{M}_{\mathrm{pt}}(E) \quad \textit{is measurable.}$$

(ii) There exists a measurable function $\Phi : \mathscr{M}_{\mathrm{pt}}(E) \to \mathbb{R}_+ \times \mathtt{Exc}$ such that

$$\forall H \in \mathtt{Exc},\; \forall t \in (0,\zeta_H), \quad \Phi(\mathcal{M}_{0,t}(H)) = (t,H)\;.$$

Proof. See Appendix 3.5. ■

Decomposition according to the total height. Let us fix $H \in \texttt{Exc}$. We introduce the first time that realizes the total height:

$$\tau(H) = \inf\{t \in \mathbb{R}_+ : H_t = \Gamma(H)\} . \tag{3.77}$$

For all $x \in (0, \Gamma(H))$ we also introduce the following times:

$$\tau_x^-(H) := \sup\big\{t < \tau(H) : H_t < \Gamma(H) - x\big\} \quad \text{and} \quad \tau_x^+(H) := \inf\big\{t > \tau(H) : H_t < \Gamma(H) - x\big\}. \tag{3.78}$$

Recall from (3.8) the definition of $H^{[s]}$. We then set

$$\forall\, t \in \mathbb{R}_+, \qquad H_t^{\ominus x} = H^{[\tau_x^-]}_{t\wedge(\tau_x^+ - \tau_x^-)} \quad \text{and} \quad H_t^{\oplus x} = H^{[\tau_x^+]}_{t\wedge(\zeta-(\tau_x^+ - \tau_x^-))} \tag{3.79}$$

where we denote $\tau_x^- := \tau_x^-(H)$, $\tau_x^+ := \tau_x^+(H)$ and $\zeta := \zeta_H$ to simplify notation. See Figure 3.2.

Let us interpret $H^{\ominus x}$ and $H^{\oplus x}$ in terms of $\mathcal{T}_H$. To that end, we recall that $p_H : [0, \zeta] \to \mathcal{T}_H$ stands for the canonical projection and we set $\gamma := p_H(\tau(H))$. We first note that $d_H(\tau_x^-, \tau_x^+) = 0$. Then we set $\gamma(x) := p_H(\tau_x^-) = p_H(\tau_x^+)$ that is the unique point of $[\![\rho, \gamma]\!]$ such that $x = d(\gamma, \gamma(x))$ and thus, $d(\rho, \gamma(x)) = \Gamma(H) - x$. We denote by $\mathcal{T}^o$ the connected component of $\mathcal{T}_H \backslash \{\gamma(x)\}$ that contains the root ρ and we set

$$\mathcal{T}^{-x} = \mathcal{T}_H \backslash \mathcal{T}^o \quad \text{and} \quad \mathcal{T}^{+x} = \{\gamma(x)\} \cup \mathcal{T}^o .$$

Thus $(\mathcal{T}^{-x}, d, \gamma(x))$ is coded by $H^{\ominus x}$ and $(\mathcal{T}^{+x}, d, \gamma(x))$ is coded by $H^{\oplus x}$. See Figure 3.2.

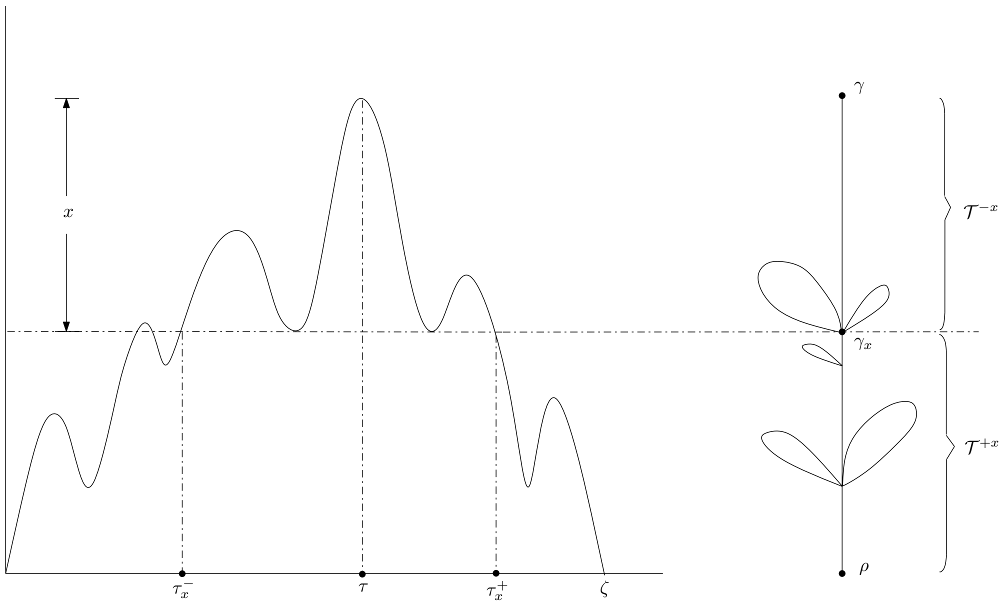

Figure 3.2 – *The left hand side figure illustrate the decomposition of H into $H^{\ominus x}$ and $H^{\oplus x}$. The right hand side figure represent this decomposition in terms of the tree coded by H.*

Recall from (3.8) the spinal decomposition at time $\tau(H)$. We shall use the following notation:

$$\mathcal{M}_{0,\tau(H)}(H) = \sum_{a \in \mathcal{J}_{0,\tau(H)}} \delta_{(a, \overleftarrow{H}^a, \overrightarrow{H}^a)}$$

that is a measure on $[0, \Gamma(H)] \times \texttt{Exc}$. Let us first make the following remark.

Remark 28. *Let $x \in (0, \Gamma(H))$ and recall the notation $\gamma(x) = p_H(\tau_x^-(H)) = p_H(\tau_x^+(H))$. Observe that if $x \notin \mathcal{J}_{0,\tau(H)}$, then $H_t > \Gamma(H) - x$, for all $t \in (\tau_x^-(H), \tau_x^+(H))$ and thus, $\tau_x^-(H)$, $\tau_x^+(H)$ are the only time $t \in [0, \zeta_H]$ such that $p_H(t) = \gamma(x)$, which implies that $\gamma(x)$ is not a branching point of $\mathcal{T}_H$: since it is not a leaf, it has to be a simple point of $\mathcal{T}_H$.* □

For all $x \in (0, \Gamma(H))$, we next introduce the following restriction of $\mathcal{M}_{0,\tau(H)}(H)$:

$$\mathcal{M}^{-x}_{0,\tau(H)}(H) = \sum_{a \in \mathcal{J}_{0,\tau(H)} \cap [0,x]} \delta_{(a, \overleftarrow{H}^a, \overrightarrow{H}^a)} \quad \text{and} \quad \mathcal{M}^{+x}_{0,\tau(H)}(H) = \sum_{a \in \mathcal{J}_{0,\tau(H)} \cap (x, \Gamma(H)]} \delta_{(a, \overleftarrow{H}^a, \overrightarrow{H}^a)} , \tag{3.80}$$

so that $\mathcal{M}_{0,\tau(H)}(H) = \mathcal{M}^{-x}_{0,\tau(H)}(H) + \mathcal{M}^{+x}_{0,\tau(H)}(H)$. Observe that

$$\tau(H) = \tau^-_x(H) + \tau(H^{\ominus x}) \quad \text{and} \quad \mathcal{M}_{0,\tau(H^{\ominus x})}(H^{\ominus x}) = \mathcal{M}^{-x}_{0,\tau(H)}(H) . \tag{3.81}$$

For all $H' \in \mathtt{Exc}$, we denote by $\Lambda(H') := (H'_{(\zeta_{H'} - t)_+})_{t \geq 0}$ the function H' that is reversed at its lifetime. We easily check that $\Lambda : \mathtt{Exc} \to \mathtt{Exc}$ is measurable and we also set:

$$\Lambda\big(\mathcal{M}^{+x}_{0,\tau(H)}(H)\big) = \sum_{a \in \mathcal{J}_{0,\tau(H)} \cap (x, \Gamma(H)]} \delta_{(\Gamma(H) - a \, , \, \Lambda(\overrightarrow{H}^a) \, , \, \Lambda(\overleftarrow{H}^a))} .$$

It is easy to check first that $\Lambda\big(\mathcal{M}^{+x}_{0,\tau(H)}(H)\big)$ is a measurable function of $\mathcal{M}^{+x}_{0,\tau(H)}(H)$ and next that

$$\mathcal{M}_{0, \zeta_H - \tau^+_x(H)}(H^{\oplus x}) = \Lambda\big(\mathcal{M}^{+x}_{0,\tau(H)}(H)\big) . \tag{3.82}$$

This combined with (3.81) and Lemma 3.9 immediately implies the following lemma.

Lemma 3.10. *There exists two measurable functions $\Phi, \Phi' : \mathscr{M}_{\mathrm{pt}}(E) \to \mathbb{R}_+ \times \mathtt{Exc}$ such that*

$$\forall H \in \mathtt{Exc}, \; \forall x \in (0, \Gamma(H)), \quad \Phi\big(\mathcal{M}^{-x}_{0,\tau(H)}(H)\big) = \big(\tau(H) - \tau^-_x(H) \, , \, H^{\ominus x}\big)$$
$$\text{and} \quad \Phi'\big(\mathcal{M}^{+x}_{0,\tau(H)}(H)\big) = \big(\zeta_H - \tau^+_x(H) \, , \, H^{\oplus x}\big) ,$$

where $\tau(H)$ is defined by (3.77), $\tau^-_x(H)$ and $\tau^+_x(H)$ by (3.78), $H^{\ominus x}$ and $H^{\oplus x}$ by (3.79) and $\mathcal{M}^{-x}_{0,\tau(H)}(H)$ and $\mathcal{M}^{+x}_{0,\tau(H)}(H)$ by (3.80).

3.2.2 Proofs of Theorem 3.1 and of Theorem 3.2.

As already mentioned, Abraham & Delmas in [3] make sense of the conditioned law $\mathbf{N}(\,\cdot\,|\,\Gamma = r)$: namely they prove that $\mathbf{N}(\,\cdot\,|\,\Gamma = r)$-a.s. $\Gamma = r$, that $r \mapsto \mathbf{N}(\,\cdot\,|\,\Gamma = r)$ is weakly continuous on $\mathbb{C}(\mathbb{R}_+, \mathbb{R}_+)$ and that (3.26) holds true. Recall from (3.30) and (3.34) the short-hand notations

$$\forall r, b, y \in (0, \infty), \quad \mathbf{N}^\Gamma_r = \mathbf{N}(\,\cdot\,|\,\Gamma = r), \quad \mathbf{N}_b = \mathbf{N}\big(\,\cdot\, \cap \{\Gamma \leq b\}\big) \quad \text{and} \quad \mathbf{P}^y_b = \mathbf{P}^y\big(\,\cdot\, \cap \{\Gamma \leq b\}\big), \tag{3.83}$$

where we recall from (3.16) the notation $\mathbf{P}^y$. Also recall from (3.23) that $\mathbf{N}^\Gamma_r$-a.s. there exists a unique $\tau \in [0, \zeta]$ such that $H_\tau = \Gamma$. Recall from (3.11) that $\mathcal{M}_{0,\tau}(H)$ gives the excursions coding the trees grafted on $[\![\rho, p(\tau)]\!]$ listed according to their distance of their grafting point from $p(\tau)$ (here $p : [0, \zeta] \to \mathcal{T}$ stands for the canonical projection). In the following lemma, we recall from Abraham & Delmas [3] the following Poisson decomposition of H under $\mathbf{N}^\Gamma_r$ at its maximum, which extends William's decomposition that corresponds to the Brownian case.

Lemma 3.11 (Abraham & Delmas [3])**.** *Let Ψ be a branching mechanism of the form (3.13) that satisfies (3.14). We keep the previous notation. Let $r \in (0, \infty)$. Then, under $\mathbf{N}^\Gamma_r$,*

$$\mathcal{M}_{0,\tau}(da \, d\overleftarrow{H} \, d\overrightarrow{H}) = \sum_{j \in \mathcal{J}_{0,\tau}} \delta_{(a, \overleftarrow{H}^a, \overrightarrow{H}^a)} \tag{3.84}$$

is Poisson point process on $[0, r] \times \mathbb{C}(\mathbb{R}_+, \mathbb{R}_+)^2$ whose intensity is

$$\mathbf{n}_r(da \, d\overleftarrow{H} \, d\overrightarrow{H}) := \beta \mathbf{1}_{[0,r]}(a) da \left(\delta_{\mathbf{0}}(d\overleftarrow{H}) \mathbf{N}_a(d\overrightarrow{H}) + \mathbf{N}_a(d\overleftarrow{H}) \delta_{\mathbf{0}}(d\overrightarrow{H})\right)$$
$$+ \; \mathbf{1}_{[0,r]}(a) da \int_{(0,\infty)} \pi(dz) \int_0^z dx \; \mathbf{P}^x_a(d\overleftarrow{H}) \, \mathbf{P}^{z-x}_a(d\overrightarrow{H}), \tag{3.85}$$

where β and π are defined in (3.13) and where $\mathbf{0}$ stands for the null function.

We first discuss several consequences of Lemma 3.11. To that end, we set

$$\boldsymbol{\nu}_{r,a}(d\overleftarrow{H}\,d\overrightarrow{H}) = \beta\delta_{\mathbf{0}}(d\overleftarrow{H})\mathbf{N}_a(d\overrightarrow{H}) + \beta\mathbf{N}_a(d\overleftarrow{H})\delta_{\mathbf{0}}(d\overrightarrow{H}) + \int_{(0,\infty)}\pi(dz)\int_0^z dx\; \mathbf{P}_a^x(d\overleftarrow{H})\,\mathbf{P}_a^{z-x}(d\overrightarrow{H}),$$

so that $\mathbf{n}_r(da\,d\overleftarrow{H}\,d\overrightarrow{H}) = \mathbf{1}_{[0,r]}(a)da\,\boldsymbol{\nu}_{r,a}(d\overleftarrow{H}\,d\overrightarrow{H})$. Denote by $\langle \boldsymbol{\nu}_{r,a}\rangle$ the total mass of $\boldsymbol{\nu}_{r,a}$. We claim that $\langle \boldsymbol{\nu}_{r,a}\rangle = \infty$.

Indeed, first recall that $\mathbf{N}$ is an infinite measure. Since $\mathbf{N}(\Gamma > a) < \infty$ (by (3.24)), $\mathbf{N}_a$ is also an infinite measure. Thus, if $\beta > 0$, $\langle \boldsymbol{\nu}_{r,a}\rangle = \infty$. Suppose now that $\beta = 0$. Then by (3.25), we get $\langle \boldsymbol{\nu}_{r,a}\rangle = \int_{(0,\infty)} \pi(dz) z e^{-zv(a)} = \infty$, since $\int_{(0,\infty)} z\,\pi(dz) = \infty$, by (3.18).

Therefore, standard results on Poisson point measures entail that $\mathbf{N}_r^\Gamma$-a.s. the closure of $\mathcal{J}_{0,\tau}$ is $[0,r]$. Recall from (3.73) the definition of $\mathtt{Exc}$ and recall from (3.17) and from (3.22) that $\mathbf{P}^x$ and $\mathbf{N}$ are supported by $\mathtt{Exc}$. Thus for all $a \in (0,r)$, $\mathbf{P}_a^x$ and $\mathbf{N}_a$ are also supported by $\mathtt{Exc}$. This entails that $\mathbf{N}_r^\Gamma$-a.s. satisfies (3.76), namely $\mathbf{N}_r^\Gamma$-a.s. $\mathcal{M}_{0,\tau} \in \mathscr{M}_{\mathrm{pt}}(E)$, where the set of point measures $\mathscr{M}_{\mathrm{pt}}(E)$ is defined in Definition 3.2.1. Then by Lemma 3.9, $\mathbf{N}_r^\Gamma$-a.s. $\Phi(\mathcal{M}_{0,\tau}) = (\tau, H)$, where Φ is a measurable function from $\mathscr{M}_{\mathrm{pt}}(E)$ to $\mathbb{R}_+ \times \mathtt{Exc}$. Thus, $\mathbf{N}_r^\Gamma$-a.s.

$$\mathbf{N}_r^\Gamma\text{-a.s.} \quad H \in \mathtt{Exc}\ . \tag{3.86}$$

Recall that $\Lambda : \mathtt{Exc} \to \mathtt{Exc}$, its the functional that reverses excursions at their lifetime: namely for all $H \in \mathtt{Exc}$, we denote by $\Lambda(H) = (H_{(\zeta_H - t)_+})_{t\geq 0}$. We recall from Corollary 3.1.6 [51] that H and $\Lambda(H)$ have the same distribution under $\mathbf{N}$. This also implies that H and $\Lambda(H)$ have the same law under $\mathbf{P}^x$. We next claim that for all $r \in (0,\infty)$,

$$H \text{ and } \Lambda(H) \text{ have the same law under } \mathbf{N}_r^\Gamma. \tag{3.87}$$

Indeed, recall the notation (3.84) for $\mathcal{M}_{0,\tau}$ and observe that

$$\mathcal{M}_{0,\tau(\Lambda(H))}(\Lambda(H)) = \sum_{a\in\mathcal{J}_{,\tau}} \delta_{(a,\Lambda(\overrightarrow{H}^a),\Lambda(\overleftarrow{H}^a))}.$$

Since $\mathbf{N}$ and $\mathbf{P}^x$ are Λ-invariant, so are $\mathbf{N}_a$ and $\mathbf{P}_a^x$ and we easily see from Lemma 3.11 that under $\mathbf{N}_r^\Gamma$, $\mathcal{M}_{0,\tau(\Lambda(H))}(\Lambda(H))$ and $\mathcal{M}_{0,\tau}$ have the same law. This implies by Lemma 3.9 that under $\mathbf{N}_r^\Gamma$, $\Phi(\mathcal{M}_{0,\tau(\Lambda(H))}(\Lambda(H))) = (\zeta - \tau, \Lambda(H))$ and $\Phi(\mathcal{M}_{0,\tau}) = (\tau, H)$ have the same law wich implies (3.87).

Recall from (3.77) the definition of $\tau(H)$, from (3.78) that of $\tau_x^-(H)$ and $\tau_x^+(H)$, from (3.79) that of $H^{\ominus x}$ and $H^{\oplus x}$, and from (3.80) that of $\mathcal{M}_{0,\tau(H)}^{-x}(H)$ and $\mathcal{M}_{0,\tau(H)}^{+x}(H)$. To simplify notation we simply write τ, τ_x^-, τ_x^+, $\mathcal{M}_{0,\tau}^{-x}$ and $\mathcal{M}_{0,\tau}^{+x}$. We then prove the following lemma.

Lemma 3.12. *We keep the same assumptions as in Lemma 3.11 and the notation therein. Let $x \in (0,r)$. Then, the following holds true.*

(i) Under $\mathbf{N}_r^\Gamma$, $\mathcal{M}_{0,\tau}^{-x}$ and $\mathcal{M}_{0,\tau}^{+x}$ are independent Poisson point measures.

(ii) $\mathbf{N}_r^\Gamma$-a.s. $x \notin \mathcal{J}_{0,\tau}$.

(iii) $\mathcal{M}_{0,\tau}^{-x}$ under $\mathbf{N}_r^\Gamma$ has the same law as $\mathcal{M}_{0,\tau}$ under $\mathbf{N}_x^\Gamma$. Thus the law of $H^{\ominus x}$ under $\mathbf{N}_r^\Gamma$ is $\mathbf{N}_x^\Gamma$.

Proof. Point (i) is a consequence of Lemma 3.11 and of basic results on Poisson point measures. Moreover, $\mathcal{M}_{0,\tau}^{-x}$ under $\mathbf{N}_r^\Gamma$ has intensity $\mathbf{1}_{[0,x]}(a)da\boldsymbol{\nu}_{r,a}(d\overleftarrow{H}\,d\overrightarrow{H})$ which is equal to $\mathbf{n}_x$. This implies that $\mathcal{M}_{0,\tau}^{-x}$ under $\mathbf{N}_r^\Gamma$ has the same law as $\mathcal{M}_{0,\tau}$ under $\mathbf{N}_x^\Gamma$. By Lemma 3.9 and Lemma 3.10, it implies that

$$(\tau - \tau_x^-, H^\ominus) = \Phi(\mathcal{M}_{0,\tau}^{-x}) \quad \text{under} \quad \mathbf{N}_r^\Gamma \quad \overset{\text{law}}{=} \quad (\tau, H) = \Phi(\mathcal{M}_{0,\tau}) \quad \text{under} \quad \mathbf{N}_x^\Gamma\ ,$$

which entails (iii). Since the intensity measure $\mathbf{n}_r(da\,d\overleftarrow{H}\,d\overrightarrow{H})$ is diffuse in the variable a, standard results on Poisson point measures entail (ii). ∎

Proof of Theorem 3.1 (i). We keep the previous notation and we set

$$\forall b\in(0,\infty),\ \forall \overleftarrow{H},\overrightarrow{H}\in \mathtt{Exc},\qquad \Delta_{b,\overleftarrow{H},\overrightarrow{H}}=b+\Gamma(\overleftarrow{H})\vee\Gamma(\overrightarrow{H})\,. \tag{3.88}$$

Recall from (3.24) and (3.25) that the distributions of Γ under $\mathbf{N}$ and under $\mathbf{P}^x$ are diffuse. Thus, for all $a\in(0,\infty)$, the distributions of Γ under $\mathbf{N}_a$ and under $\mathbf{P}^x_a$ are also diffuse. Recall the notation (3.84) for $\mathcal{M}_{0,\tau}$. Then, Lemma 3.11 combined with Lemma 3.8 implies that $\mathbf{N}^\Gamma_r$-a.s. there exists a unique $Y\in(0,r)\cap\mathcal{J}_{0,\tau}$ such that

$$D=Y+\Gamma(\overleftarrow{H}^Y)\vee\Gamma(\overrightarrow{H}^Y)=\Delta_{Y,\overleftarrow{H}^Y,\overrightarrow{H}^Y}>\sup_{a\in\mathcal{J}_{0,\tau}\backslash\{Y\}}\Delta_{a,\overleftarrow{H}^a,\overrightarrow{H}^a}\,. \tag{3.89}$$

Then either $\Gamma(\overleftarrow{H}^Y)<\Gamma(\overrightarrow{H}^Y)$ or $\Gamma(\overleftarrow{H}^Y)>\Gamma(\overrightarrow{H}^Y)$. Let us us consider these two cases.

• If $\Gamma(\overleftarrow{H}^Y)<\Gamma(\overrightarrow{H}^Y)$ then by (3.23) and (3.25) there exists a unique point t_* such that $\overrightarrow{H}^Y_{t_*}=\Gamma(\overrightarrow{H}^Y)$. This proves Theorem 3.1 (i) in this case under $\mathbf{N}^\Gamma_r$ and we have $\tau_0=\tau$ and

$$\tau_1=\tau+t_*+\sum_{a\in\mathcal{J}_{0,\tau}\cap[0,Y)}\zeta_{\overrightarrow{H}^a}\,.$$

• If $\Gamma(\overleftarrow{H}^Y)>\Gamma(\overrightarrow{H}^Y)$ then by (3.23) and (3.25) there exists a unique point t_* such that $\overleftarrow{H}^Y_{t_*}=\Gamma(\overleftarrow{H}^Y)$. This proves Theorem 3.1 (i) in this case under $\mathbf{N}^\Gamma_r$ and we have $\tau_1=\tau$ and

$$\tau_0=t_*+\sum_{a\in\mathcal{J}_{0,\tau}\cap(Y,r]}\zeta_{\overleftarrow{H}^a}\,.$$

Theorem 3.1 (i) is then proved under $\mathbf{N}^\Gamma_r$, for all $r\in(0,\infty)$, which implies Theorem 3.1 (i) (under $\mathbf{N}$) by (3.26). $\square$

Proof Theorem 3.1 (ii). Recall from (3.80) the notation $\mathcal{M}^{-x}_{0,\tau}$ and $\mathcal{M}^{+x}_{0,\tau}$. We shall use the following lemma.

Lemma 3.13. *We keep the same assumptions as in Lemma 3.11 and the notation therein. Recall from Definition 3.2.1 the notation $\mathscr{M}_{\mathrm{pt}}(E)$. Then, for all $r\in(0,\infty)$ and for all measurable functions $G_1,G_2:\mathscr{M}_{\mathrm{pt}}(E)\to\mathbb{R}_+$,*

$$\mathbf{N}^\Gamma_r\Big[\mathbf{1}_{\{\tau=\tau_0\}}G_1\big(\mathcal{M}^{-\frac12 D}_{0,\tau}\big)G_2\big(\mathcal{M}^{+\frac12 D}_{0,\tau}\big)\Big]=\mathbf{N}^\Gamma_r\Big[\mathbf{1}_{\{\tau=\tau_0\}}\mathbf{N}^\Gamma_{\frac12 D}[G_1(\mathcal{M}_{0,\tau})]G_2\big(\mathcal{M}^{+\frac12 D}_{0,\tau}\big)\Big]\,,$$

with a similar statements where τ_0 is replaced by τ_1. Moreover, by (3.26) a similar statement holds true under $\mathbf{N}$.

Before proving this lemma, we first complete the proof of Theorem 3.1. Recall from the notation (3.84) and from (3.80) that

$$\mathcal{M}_{0,\tau}=\sum_{j\in\mathcal{J}_{0,\tau}}\delta_{(a,\overleftarrow{H}^a,\overrightarrow{H}^a)}\quad\text{and}\quad\mathcal{M}^{-\frac12 D}_{0,\tau}=\sum_{j\in\mathcal{J}_{0,\tau}\cap[0,\frac12 D]}\delta_{(a,\overleftarrow{H}^a,\overrightarrow{H}^a)}\,.$$

By Lemma 3.13, we then get

$$\mathbf{N}\Big(\tfrac12 D\in\mathcal{J}_{0,\tau}\Big)=\mathbf{N}\Big[\mathbf{N}^\Gamma_{\frac12 D}\Big(\tfrac12 D\in\mathcal{J}_{0,\tau}\Big)\Big]=0$$

because for any $b\in(0,\infty)$, Lemma 3.11 asserts that under $\mathbf{N}^\Gamma_b$, $\mathcal{M}_{0,\tau}$ is a Poisson point measure with intensity $\mathbf{n}_b$, which implies that $\mathbf{N}^\Gamma_b$-a.s. $b\notin\mathcal{J}_{0,\tau}$. We next use Remark 28 with $x=\frac12 D$ that asserts that

$$\tau^-_{\mathrm{mid}}:=\tau^-_{\frac12 D}\quad\text{and}\quad\tau^+_{\mathrm{mid}}:=\tau^+_{\frac12 D} \tag{3.90}$$

are the only times $t\in[0,\zeta]$, such that $d(p(\tau_1),p(t))=\frac12 D$, which completes the proof of Theorem 3.1 (ii). $\square$

Proof Theorem 3.1 (iii)**.** Let $r, y \in (0, \infty)$ be such that $\frac{1}{2}y < r < y$. We first work under $\mathbf{N}_r^\Gamma$. Recall from (3.84) the notation for $\mathcal{M}_{0,\tau}$ and recall notation (3.88). Then (3.89) combined with Lemma 3.11 that asserts that under $\mathbf{N}_r^\Gamma$, $\mathcal{M}_{0,\tau}$ is a Poisson point measure with intensity $\mathbf{n}_r$, we get

$$\mathbf{N}_r^\Gamma\big(D \le y\big) = \mathbf{N}_r^\Gamma\big(\sup\{\Delta_{a,\overleftarrow{H}^a,\overrightarrow{H}^a}\, ;\, a \in \mathcal{J}_{0,\tau}\} \le y\big) = \exp\Big(-\int \mathbf{n}_r(da\, d\overleftarrow{H}\, d\overrightarrow{H})\mathbf{1}_{\{\Delta_{a,\overleftarrow{H}^a,\overrightarrow{H}^a} > y\}}\Big), \tag{3.91}$$

where $\mathbf{n}_r$ is given by (3.85). Recall from (3.24) that $\mathbf{N}(\Gamma > t) = v(t)$ and from (3.25) that $\mathbf{P}^x(\Gamma \le t) = e^{-xv(t)}$. Thus,

$$\begin{aligned}\int \mathbf{n}_r(da\, d\overleftarrow{H}\, d\overrightarrow{H})\mathbf{1}_{\{\Delta_{a,\overleftarrow{H}^a,\overrightarrow{H}^a} > y\}} &= 2\beta \int_0^r da\, \mathbf{N}(y - a < \Gamma \le a) \\ &+ \int_0^r da \int_{(0,\infty)} \pi(dz) \int_0^z dx \int \mathbf{P}_a^x(d\overleftarrow{H}) \int \mathbf{P}_a^{z-x}(d\overrightarrow{H})\big(1 - \mathbf{1}_{\{\Gamma(\overleftarrow{H}) \le y-a\}} \mathbf{1}_{\{\Gamma(\overrightarrow{H}) \le y-a\}}\big).\end{aligned}$$

Recall from (3.24) that $\mathbf{N}(\Gamma > t) = v(t)$. Thus, if $a < \frac{1}{2}y$, then $\mathbf{N}(y - a < \Gamma \le a) = 0$ and if $a > \frac{1}{2}y$, then $\mathbf{N}(y - a < \Gamma \le a) = v(y - a) - v(a)$. Next recall from (3.25) that $\mathbf{P}^x(\Gamma \le t) = e^{-xv(t)}$, which implies that the total mass of $\mathbf{P}_a^x$ is $\mathbf{P}^x(\Gamma \le a) = \exp(-xv(a))$. Also observe that $\mathbf{P}_a^x(\Gamma \le y - a) = \mathbf{P}^x(\Gamma \le a \wedge (y - a)) = \exp(-xv(a \wedge (y - a))$. Thus

$$\int \mathbf{P}_a^x(d\overleftarrow{H}) \int \mathbf{P}_a^{z-x}(d\overrightarrow{H})\big(1 - \mathbf{1}_{\{\Gamma(\overleftarrow{H}) \le y-a\}} \mathbf{1}_{\{\Gamma(\overrightarrow{H}) \le y-a\}}\big) = e^{-zv(a)} - e^{-zv(a \wedge (y-a))},$$

which is null if $a < \frac{1}{2}y$. Note that this expression does not depend on x. Thus,

$$\begin{aligned}\int \mathbf{n}_r(da\, d\overleftarrow{H}\, d\overrightarrow{H})\mathbf{1}_{\{\Delta_{a,\overleftarrow{H}^a,\overrightarrow{H}^a} > y\}} &= \int_{\frac{1}{2}y}^r da\, 2\beta\big(v(y-a) - v(a)\big) + \int_{\frac{1}{2}y}^r da \int_{(0,\infty)} \pi(dz)\, z\big(e^{-zv(a)} - e^{-zv(y-a)}\big) \\ &= \int_{\frac{1}{2}y}^r da\big(\Psi'(v(y-a)) - \Psi'(v(a))\big) = \int_{y-r}^{\frac{1}{2}y} db\, \Psi'(v(b)) - \int_{\frac{1}{2}y}^r db\, \Psi'(v(b)).\end{aligned}$$

by (3.13). Recall that v satisfies $\int_{v(b)}^\infty d\lambda/\Psi(\lambda) = b$. By the change of variable $\lambda = v(b)$, we then get

$$\begin{aligned}\int \mathbf{n}_r(da\, d\overleftarrow{H}\, d\overrightarrow{H})\mathbf{1}_{\{\Delta_{a,\overleftarrow{H}^a,\overrightarrow{H}^a} > y\}} &= \int_{v(\frac{1}{2}y)}^{v(y-r)} d\lambda\, \frac{\Psi'(\lambda)}{\Psi(\lambda)} - \int_{v(r)}^{v(\frac{1}{2}y)} d\lambda\, \frac{\Psi'(\lambda)}{\Psi(\lambda)} \\ &= \log \frac{\Psi(v(y-r))}{\Psi(v(\frac{1}{2}y))} - \log \frac{\Psi(v(\frac{1}{2}y))}{\Psi(v(r))}.\end{aligned}$$

By (3.91), we get

$$\forall r \in (0, \infty),\ \forall y \in (r, 2r), \quad \mathbf{N}_r^\Gamma\big(D \le y\big) = \frac{\Psi(v(\frac{1}{2}y))^2}{\Psi(v(r))\Psi(v(y-r))}\,. \tag{3.92}$$

Now observe that $\mathbf{N}_r^\Gamma\big(D \le y\big) = 1$, if $y \ge 2r$ and that $\mathbf{N}_r^\Gamma\big(D \le y\big) = 0$, if $y \le r$. Thus by (3.26),

$$\begin{aligned}\mathbf{N}(D \ge y) &= \int_0^\infty \mathbf{N}(\Gamma \in dr)\, \mathbf{N}_r^\Gamma(D \ge y) = \mathbf{N}(\Gamma \ge y) + \int_{\frac{1}{2}y}^y dr\, \Psi(v(r))\Big(1 - \frac{\Psi(v(\frac{1}{2}y))^2}{\Psi(v(r))\Psi(v(y-r))}\Big) \\ &= v(\tfrac{1}{2}y) - \Psi(v(\tfrac{1}{2}y))^2 \int_{\frac{1}{2}y}^y \frac{dr}{\Psi(v(y-r))} = v(\tfrac{1}{2}y) - \Psi(v(\tfrac{1}{2}y))^2 \int_{v(\frac{1}{2}y)}^\infty \frac{d\lambda}{\Psi(\lambda)^2},\end{aligned}$$

where we use the change of variable $\lambda = v(y-r)$ in the last equality. This proves (3.27) that easily entails (3.28), which completes the proof of Theorem 3.1 (iii). □

Proof of Lemma 3.13. To completes the proof of Theorem 3.1, it remains to prove Lemma 3.13 that is also the key argument to prove Theorem 3.2. We first work under $\mathbf{N}_r^\Gamma$. Recall the notation (3.84) for $\mathcal{M}_{0,\tau}$ and $\mathcal{J}_{0,\tau}$ and recall from (3.80) the following definitions (with $x = \frac{1}{2}D$),

$$\mathcal{M}_{0,\tau} = \sum_{j\in\mathcal{J}_{0,\tau}} \delta_{(a,\overleftarrow{H}^a,\overrightarrow{H}^a)}\, , \quad \mathcal{M}_{0,\tau}^{-\frac{1}{2}D} = \sum_{j\in\mathcal{J}_{0,\tau}\cap[0,\frac{1}{2}D]} \delta_{(a,\overleftarrow{H}^a,\overrightarrow{H}^a)} \quad \text{and} \quad \mathcal{M}_{0,\tau}^{+\frac{1}{2}D} = \sum_{j\in\mathcal{J}_{0,\tau}\cap(\frac{1}{2}D,r]} \delta_{(a,\overleftarrow{H}^a,\overrightarrow{H}^a)}.$$

Recall from (3.89) the definition of the random variable Y: since $\Gamma(\overleftarrow{H}^Y)\vee\Gamma(\overrightarrow{H}^Y) < Y$, we get $Y > \frac{1}{2}D$ and $(Y, \overleftarrow{H}^Y, \overrightarrow{H}^Y)$ is an atom of $\mathcal{M}_{0,\tau}^{+\frac{1}{2}D}$. This argument, combined with (3.89) and the Palm formula for Poisson point measures, implies

$$\begin{aligned}\mathbf{N}_r^\Gamma\Big[\mathbf{1}_{\{\tau=\tau_0\}}F(Y,\overleftarrow{H}^Y,\overrightarrow{H}^Y)G_1(\mathcal{M}_{0,\tau}^{-\frac{1}{2}D})G_2(\mathcal{M}_{0,\tau}^{+\frac{1}{2}D})\Big] = \\ \int \mathbf{n}_r(dy\, dH'\, dH'')\mathbf{1}_{\{\Gamma(H'')>\Gamma(H')\}}F(y,H',H'') \\ \times \mathbf{N}_r^\Gamma\Big[G_1(\mathcal{M}_{0,\tau}^{-\frac{1}{2}\Delta_{y,H',H''}})G_2\Big(\mathcal{M}_{0,\tau}^{+\frac{1}{2}\Delta_{y,H',H''}} + \delta_{(y,H',H'')}\Big)\mathbf{1}_{\{\Delta_{y,H',H''}>\sup\{\Delta_{a,\overleftarrow{H}^a,\overrightarrow{H}^a}\, ;\, a\in\mathcal{J}_{0,\tau}\}\}}\Big].\end{aligned} \tag{3.93}$$

where we recall that $\tau_0 = \tau$ iff $\Gamma(\overrightarrow{H}^Y) > \Gamma(\overleftarrow{H}^Y)$. Then observe that $\mathbf{n}_r\otimes\mathbf{N}_r^\Gamma$-a.e. for all $a \in \mathcal{J}_{0,\tau}\cap [0, \frac{1}{2}\Delta_{y,H',H''}]$, we have $\Delta_{a,\overleftarrow{H}^a,\overrightarrow{H}^a} < 2a \le \Delta_{y,H',H''}$. Thus, $\mathbf{n}_r\otimes\mathbf{N}_r^\Gamma$-a.e.

$$\mathbf{1}_{\{\Delta_{y,H',H''}>\sup\{\Delta_{a,\overleftarrow{H}^a,\overrightarrow{H}^a}\, ;\, a\in\mathcal{J}_{0,\tau}\}\}} = \mathbf{1}_{\{\Delta_{y,H',H''}>\sup\{\Delta_{a,\overleftarrow{H}^a,\overrightarrow{H}^a}\, ;\, a\in\mathcal{J}_{0,\tau}\cap(\frac{1}{2}\Delta_{y,H',H''},r]\}\}}$$

that only depends on y, H', H'' and of $\mathcal{M}_{0,\tau}^{+\frac{1}{2}\Delta_{y,H',H''}}$. By (3.93) with $F \equiv 1$ and by Lemma 3.12 (i) and (iii) with $x = \frac{1}{2}\Delta_{y,H',H''}$, we get

$$\begin{aligned}\mathbf{N}_r^\Gamma\Big[\mathbf{1}_{\{\tau=\tau_0\}}G_1(\mathcal{M}_{0,\tau}^{-\frac{1}{2}D})G_2(\mathcal{M}_{0,\tau}^{+\frac{1}{2}D})\Big] = \\ \int \mathbf{n}_r(dy\, dH'\, dH'')\mathbf{1}_{\{\Gamma(H'')>\Gamma(H')\}}\,\mathbf{N}^\Gamma_{\frac{1}{2}\Delta_{y,H',H''}}[G_1(\mathcal{M}_{0,\tau})] \\ \times \mathbf{N}_r^\Gamma\Big[G_2\Big(\mathcal{M}_{0,\tau}^{+\frac{1}{2}\Delta_{y,H',H''}} + \delta_{(y,H',H'')}\Big)\mathbf{1}_{\{\Delta_{y,H',H''}>\sup\{\Delta_{a,\overleftarrow{H}^a,\overrightarrow{H}^a}\, ;\, a\in\mathcal{J}_{0,\tau}\}\}}\Big]. \\ = \mathbf{N}_r^\Gamma\Big[\mathbf{1}_{\{\tau=\tau_0\}}\mathbf{N}^\Gamma_{\frac{1}{2}D}[G_1(\mathcal{M}_{0,\tau})]G_2(\mathcal{M}_{0,\tau}^{+\frac{1}{2}D})\Big],\end{aligned}$$

which completes the proof of Lemma 3.13 when $\tau = \tau_0$ under $\mathbf{N}_r^\Gamma$. When $\tau = \tau_1$, the proof is quite similar. Then, (3.26) immediately entails the same result under $\mathbf{N}$. $\square$

Proof of Theorem 3.2 (iii)**.** Lemma 3.13 under $\mathbf{N}$ and Lemma 3.10 imply that for all measurable functions $F_1, F_2 : \mathbb{C}(\mathbb{R}_+,\mathbb{R}_+)\to\mathbb{R}_+$, $f : \mathbb{R}_+ \to \mathbb{R}_+$,

$$\mathbf{N}\Big[\mathbf{1}_{\{\tau=\tau_0\}}f(D)F_1\big(H^{\ominus\frac{1}{2}D}\big)F_2\big(H^{\oplus\frac{1}{2}D}\big)\Big] = \mathbf{N}\Big[\mathbf{1}_{\{\tau=\tau_0\}}f(D)\mathbf{N}^\Gamma_{\frac{1}{2}D}[F_1(H)]F_2\big(H^{\oplus\frac{1}{2}D}\big)\Big]\, , \tag{3.94}$$

with a similar statement with $\tau = \tau_1$. To simplify notation, we next set

$$H^\ominus := H^{\ominus\frac{1}{2}D} \quad \text{and} \quad H^\oplus := H^{\oplus\frac{1}{2}D}\, .$$

By adding (3.94) with the analogous equality with $\tau = \tau_1$, we get

$$\mathbf{N}\Big[f(D)F_1\big(H^\ominus\big)F_2\big(H^\oplus\big)\Big] = \mathbf{N}\Big[f(D)\,\mathbf{N}^\Gamma_{\frac{1}{2}D}[F_1(H)]\,F_2\big(H^\oplus\big)\Big]\, . \tag{3.95}$$

Recall from (3.90) that $\tau_{\mathrm{mid}}^- = \tau_{\frac{1}{2}D}^-$ and $\tau_{\mathrm{mid}}^+ = \tau_{\frac{1}{2}D}^+$; rewriting (3.79) with $x = \frac{1}{2}D$ yields

$$H^{\ominus} = H^{[\tau_{\mathrm{mid}}^-]}_{\cdot\wedge(\tau_{\mathrm{mid}}^+ - \tau_{\mathrm{mid}}^-)}, \quad H^{\oplus} = H^{[\tau_{\mathrm{mid}}^+]}_{\cdot\wedge(\zeta - (\tau_{\mathrm{mid}}^+ - \tau_{\mathrm{mid}}^-))} \quad \text{and thus} \quad H^{[\tau_{\mathrm{mid}}^-]} = H^{\ominus} \oplus H^{\oplus}, \tag{3.96}$$

where we recall from (3.29) that $H' \oplus H''$ stands for the concatenation of the functions H' and H''.

Let us briefly interpret $H^{\ominus}$ and $H^{\oplus}$ in terms of the tree $\mathcal{T}$. To that end, first recall that $\gamma = p(\tau)$, $\gamma_0 = p(\tau_0)$ and $\gamma_1 = p(\tau_1)$, where $p : [0, \zeta] \to \mathcal{T}$ stands for the canonical projection. Recall that γ_{mid} is the mid point of the diameter $[\![\gamma_0, \gamma_1]\!]$: namely $d(\gamma_0, \gamma_{\mathrm{mid}}) = d(\gamma_1, \gamma_{\mathrm{mid}}) = \frac{1}{2}D$. Recall from Theorem 3.1 (ii) that τ_{mid}^- and τ_{mid}^+ are the only times $t \in [0, \zeta]$ such that $p(t) = \gamma_{\mathrm{mid}}$; thus, γ_{mid} is a simple point of $\mathcal{T}$; namely, $\mathcal{T} \backslash \{\gamma_{\mathrm{mid}}\}$ has only two connected components. Denote by $\mathcal{T}^o$ the connected component containing γ: it does not contain the root; if we set $\mathcal{T}^- = \{\gamma_{\mathrm{mid}}\} \cup \mathcal{T}^o$ and $\mathcal{T}^+ = \mathcal{T} \backslash \mathcal{T}^o$, then $H^{\ominus}$ codes $(\mathcal{T}^-, d, \gamma_{\mathrm{mid}})$ and $H^{\oplus}$ codes $(\mathcal{T}^+, d, \gamma_{\mathrm{mid}})$.

We next use Proposition 2.1 from D. & Le Gall [53] that is recalled as follows.

Lemma 3.14 (D. & Le Gall [53]). *For all measurable functions $F : \mathbb{R}_+ \times \mathbb{C}(\mathbb{R}_+, \mathbb{R}_+) \to \mathbb{R}_+$ and $g : \mathbb{R}_+ \to \mathbb{R}_+$,*

$$\mathbf{N}\Big[g(\zeta)\int_0^{\zeta} dt\, F\big(t, H^{[t]}\big)\Big] = \mathbf{N}\Big[g(\zeta)\int_0^{\zeta} dt\, F(t, H)\Big] .$$

This result asserts that H is invariant under uniform re-rooting. By applying this property we first get

$$\mathbf{N}\big[\zeta F_1(H^{\ominus})F_2(H^{\oplus})\big] = \mathbf{N}\Big[\int_0^{\zeta} dt\, F_1(H^{\ominus})F_2(H^{\oplus})\Big] = \mathbf{N}\Big[\int_0^{\zeta} dt\, F_1\big((H^{[t]})^{\ominus}\big)F_2\big((H^{[t]})^{\oplus}\big)\Big] . \tag{3.97}$$

Next observe the following: if $t \in (\tau_{\mathrm{mid}}^-, \tau_{\mathrm{mid}}^+)$, then $(H^{[t]})^{\ominus} = H^{\oplus}$ and $(H^{[t]})^{\oplus} = H^{\ominus}$, and if $t \in (0, \tau_{\mathrm{mid}}^-) \cup (\tau_{\mathrm{mid}}^+, \zeta)$, then $(H^{[t]})^{\ominus} = H^{\ominus}$ and $(H^{[t]})^{\oplus} = H^{\oplus}$. Thus,

$$\begin{aligned}\int_0^{\zeta} dt\, F_1\big((H^{[t]})^{\ominus}\big)F_2\big((H^{[t]})^{\oplus}\big) &= \big(\tau_{\mathrm{mid}}^+ - \tau_{\mathrm{mid}}^-\big)F_1(H^{\oplus})F_2(H^{\ominus}) + \big(\zeta - \tau_{\mathrm{mid}}^+ + \tau_{\mathrm{mid}}^-\big)F_1(H^{\ominus})F_2(H^{\oplus}) \\ &= \zeta_{H^{\ominus}} F_1(H^{\oplus})F_2(H^{\ominus}) + \zeta_{H^{\oplus}} F_1(H^{\ominus})F_2(H^{\oplus}).\end{aligned}$$

This equality, (3.97) and (3.95) with $f \equiv 1$ imply the following:

$$\begin{aligned}\mathbf{N}\big[\zeta F_1(H^{\ominus})F_2(H^{\oplus})\big] &= \mathbf{N}\big[\zeta_{H^{\ominus}} F_1(H^{\oplus})F_2(H^{\ominus})\big] + \mathbf{N}\big[\zeta_{H^{\oplus}} F_1(H^{\ominus})F_2(H^{\oplus})\big] \\ &= \mathbf{N}\Big[\mathbf{N}^{\Gamma}_{\frac{1}{2}D}\big[\zeta F_2(H)\big]\; F_1(H^{\oplus})\Big] + \mathbf{N}\Big[\mathbf{N}^{\Gamma}_{\frac{1}{2}D}[F_1(H)]\; \zeta_{H^{\oplus}} F_2\big(H^{\oplus}\big)\Big]\end{aligned} \tag{3.98}$$

Next observe that $\zeta_{H^{\ominus}} + \zeta_{H^{\oplus}} = \zeta$. Thus, by (3.95) we also get

$$\begin{aligned}\mathbf{N}\big[\zeta F_1(H^{\ominus})F_2(H^{\oplus})\big] &= \mathbf{N}\big[\zeta_{H^{\ominus}} F_1(H^{\ominus})F_2(H^{\oplus})\big] + \mathbf{N}\big[\zeta_{H^{\oplus}} F_1(H^{\ominus})F_2(H^{\oplus})\big] \\ &= \mathbf{N}\Big[\mathbf{N}^{\Gamma}_{\frac{1}{2}D}\big[\zeta F_1(H)\big]\; F_2(H^{\oplus})\Big] + \mathbf{N}\Big[\mathbf{N}^{\Gamma}_{\frac{1}{2}D}[F_1(H)]\; \zeta_{H^{\oplus}} F_2\big(H^{\oplus}\big)\Big]\end{aligned} \tag{3.99}$$

Then by (3.98) and (3.99), we get $\mathbf{N}\big[\mathbf{N}^{\Gamma}_{\frac{1}{2}D}\big[\zeta F_1(H)\big]\; F_2(H^{\oplus})\big] = \mathbf{N}\big[\mathbf{N}^{\Gamma}_{\frac{1}{2}D}\big[\zeta F_2(H)\big]\; F_1(H^{\oplus})\big]$. Since the total height of $H^{\ominus}$ and $H^{\oplus}$ is $\frac{1}{2}D$, for all measurable functions $F_1, F_2 : \mathbb{C}(\mathbb{R}_+, \mathbb{R}_+) \to \mathbb{R}_+$, $f : \mathbb{R}_+ \to \mathbb{R}_+$, we get

$$\mathbf{N}\Big[f(D)\, \mathbf{N}^{\Gamma}_{\frac{1}{2}D}\big[\zeta F_1(H)\big]\; F_2(H^{\oplus})\Big] = \mathbf{N}\Big[f(D)\, \mathbf{N}^{\Gamma}_{\frac{1}{2}D}\big[\zeta F_2(H)\big]\; F_1(H^{\oplus})\Big]. \tag{3.100}$$

By taking in (3.100) $F_1 \equiv 1$ and by substituting $f(D)$ with $f(D)/\mathbf{N}^{\Gamma}_{\frac{1}{2}D}[\,\zeta\,]$, we get

$$\mathbf{N}\big[f(D)\, F_2(H^{\oplus})\big] = \mathbf{N}\Big[f(D)\, \mathbf{N}^{\Gamma}_{\frac{1}{2}D}\big[\zeta F_2(H)\big] / \mathbf{N}^{\Gamma}_{\frac{1}{2}D}[\,\zeta\,]\Big],$$

and by (3.95), it entails

$$\mathbf{N}\big[f(D)\,F_1(H^{\ominus})F_2(H^{\oplus})\big] = \mathbf{N}\Big[f(D)\,\mathbf{N}^{\Gamma}_{\frac{1}{2}D}\big[F_1(H)\big]\,\mathbf{N}^{\Gamma}_{\frac{1}{2}D}\big[\zeta F_2(H)\big]/\mathbf{N}^{\Gamma}_{\frac{1}{2}D}\big[\zeta\big]\Big]\,. \tag{3.101}$$

Recall from (3.96) that $H^{[\tau^-_{\mathrm{mid}}]} = H^{\ominus} \oplus H^{\oplus}$. Then, (3.101) implies for all measurable functions $F : \mathbb{C}(\mathbb{R}_+, \mathbb{R}_+) \to \mathbb{R}_+$, $f : \mathbb{R}_+ \to \mathbb{R}_+$, that

$$\begin{aligned}\mathbf{N}\big[f(D)F\big(H^{[\tau^-_{\mathrm{mid}}]}\big)\big] = \\ \int_0^\infty \mathbf{N}(D\in dr)\frac{f(r)}{\mathbf{N}^{\Gamma}_{r/2}[\zeta]}\iint_{\mathbb{C}(\mathbb{R}_+,\mathbb{R}_+)^2}\mathbf{N}^{\Gamma}_{r/2}(dH)\mathbf{N}^{\Gamma}_{r/2}(dH')\;\zeta_{H'}F\big(H\oplus H'\big)\,,\end{aligned} \tag{3.102}$$

which implies Theorem 3.2 (iii) as soon as one makes sense of $\mathbf{N}(\,\cdot\,|\,D=r)$. □

Proof of Theorem 3.2 (ii)**.** Recall that $\Lambda:\mathtt{Exc}\to\mathtt{Exc}$ is the functional that reverses excursions at their lifetime: namely for all $H\in\mathtt{Exc}$, $\Lambda(H) = (H_{(\zeta_H-t)_+})_{t\geq 0}$. Recall from (3.87) that for all $r\in(0,\infty)$, H and $\Lambda(H)$ have the same law under $\mathbf{N}^{\Gamma}_r$, which entails the following by (3.101):

$$\big(\Lambda(H^{\ominus}),\Lambda(H^{\oplus})\big) \text{ and } (H^{\ominus},H^{\oplus}) \text{ have the same distribution under } \mathbf{N}. \tag{3.103}$$

Next, observe that $D(\Lambda(H))=D$, $\tau(\Lambda(H))=\zeta-\tau$, $\tau_0(\Lambda(H))=\zeta-\tau_1$ and $\tau_1(\Lambda(H))=\zeta-\tau_0$. Moreover, $(\Lambda(H))^{\ominus} = \Lambda(H^{\ominus})$ and $(\Lambda(H))^{\oplus} = \Lambda(H^{\oplus})$. This combined with (3.103) and (3.101) implies that

$$\begin{aligned}\tfrac{1}{2}\mathbf{N}\big[f(D)F_1(H^{\ominus})F_2(H^{\oplus})\big] &= \mathbf{N}\big[\mathbf{1}_{\{\tau=\tau_0\}}f(D)F_1(H^{\ominus})F_2(H^{\oplus})\big] \\ &= \mathbf{N}\big[\mathbf{1}_{\{\tau=\tau_1\}}f(D)F_1(H^{\ominus})F_2(H^{\oplus})\big]\,.\end{aligned} \tag{3.104}$$

We then define

$$\tau^*:=\tau^-_{\mathrm{mid}} \text{ if } \tau=\tau_0 \quad\text{and}\quad \tau^*:=\tau^+_{\mathrm{mid}} \text{ if } \tau=\tau_1.$$

By (3.96), we get

$$H^{[\tau^*]} = H^{\ominus}\oplus H^{\oplus} \text{ on } \{\tau=\tau_0\} \quad\text{and}\quad H^{[\tau^*]} = H^{\oplus}\oplus H^{\ominus} \text{ on } \{\tau=\tau_1\}.$$

This, combined with (3.104) and (3.101) entails

$$\begin{aligned}\mathbf{N}\big[f(D)F\big(H^{[\tau^*]}\big)\big] = \\ \int_0^\infty \mathbf{N}(D\in dr)\frac{f(r)}{2\mathbf{N}^{\Gamma}_{r/2}[\zeta]}\iint_{\mathbb{C}(\mathbb{R}_+,\mathbb{R}_+)^2}\mathbf{N}^{\Gamma}_{r/2}(dH)\,\mathbf{N}^{\Gamma}_{r/2}(dH')\;(\zeta_H+\zeta_{H'})F\big(H\oplus H'\big)\,.\end{aligned} \tag{3.105}$$

Recall from (3.31) the definition of the law $\mathbf{Q}_r$. Since $r\mapsto\mathbf{N}^{\Gamma}_r$ is weakly continuous, it is easy to check that $r\mapsto\mathbf{Q}_r$ is also weakly continuous. Then observe that $\mathbf{Q}_r[\zeta] = 2\mathbf{N}^{\Gamma}_{r/2}[\zeta]$. Therefore (3.105) can be rewritten as

$$\mathbf{N}\big[f(D)F\big(H^{[\tau^*]}\big)\big] = \int_0^\infty \mathbf{N}(D\in dr)\,f(r)\,\mathbf{Q}_r\big[\zeta F(H)\big]/\mathbf{Q}_r[\zeta]\,. \tag{3.106}$$

Next observe that for all $t\in[0,\zeta]$, $(H^{[\tau^*]})^{[t]} = H^{[\tau^*+t]}$ and that $D(H^{[t]})=D$. Thus, (3.106) implies

$$\begin{aligned}\int_0^\infty \mathbf{N}(D\in dr)\,f(r)\mathbf{Q}_r\Big[\zeta\int_0^\zeta dt\,F\big(H^{[t]}\big)\Big]/\mathbf{Q}_r[\zeta] &= \mathbf{N}\Big[f(D)\int_0^\zeta dt\,F\big(H^{[\tau^*+t]}\big)\Big] \\ &= \mathbf{N}\Big[\int_0^\zeta dt\,f\big(D(H^{[t]})\big)F\big(H^{[t]}\big)\Big] \\ &= \mathbf{N}\big[\zeta f(D)F(H)\big]\,,\end{aligned}$$

where we have use Lemma 3.14 in the last line. This proves (3.32) in Theorem 3.2 (ii).

Proof of Theorem 3.2 (i) **and** (iv). The rest of the proof is now easy: we fix $r \in (0,\infty)$ and we denote by $\Pi_r(dH'dH'')$ the the product law $\mathbf{N}^{\Gamma}_{r/2}(dH')\mathbf{N}^{\Gamma}_{r/2}(dH'')$; we then set $H = H' \oplus H''$. Thus, by definition, H under Π_r has law $\mathbf{Q}_r$. Observe that if $t \neq \tau(H')$ (resp. $t \neq \tau(H'')$) then $H'_t < r/2$ (resp. $H''_t < r/2$). Note that if $s \in [0, \zeta_{H'}]$ and $t \in [\zeta_{H'}, \zeta_{H''}]$, then $\inf_{[s,t]} H = 0$ and $d_H(s,t) = H'_s + H''_{t-\zeta_{H'}}$. This easily entails that Π_r-a.s. $D(H) = r$ and that $\tau(H')$ and $\zeta_{H'} + \tau(H'')$ are the two only times $s<t$ such that $d_H(s,t) = D(H)$, which completes the proof of 3.2 (i).

The fact that $\mathbf{Q}_r$-a.s. $D = r$, combined with (3.32) and with the fact that $r \mapsto \mathbf{Q}_r$ is weakly continuous, allows to make sense of $\mathbf{N}(\,\cdot\,|\, D = r)$ that is a regular version of the conditional distribution of $\mathbf{N}$ knowing that $D = r$. Moreover, (3.32) entails (3.36) for all $r \in (0,\infty)$. Furthermore (3.102) entails (3.33) that was the last point to clear in the Theorem 3.2 (iii), as already mentioned.

It remains to prove Theorem 3.2 (iv). We keep the previous notations and we introduce the following:

$$\mathcal{M}_{0,\tau(H')}(H') = \sum_{a \in \mathcal{J}_{0,\tau'}} \delta_{(a, \overleftarrow{H}^a, \overrightarrow{H}^a)} \quad \text{and} \quad \mathcal{M}_{0,\tau(H'')}(H'') = \sum_{a \in \mathcal{J}_{0,\tau''}} \delta_{(a, \overleftarrow{H}^a, \overrightarrow{H}^a)} \, ,$$

that are under Π_r independent Poisson point measures with the same intensity $\mathbf{n}_{r/2}$, by Lemma 3.11. We then set $\tau_0(H) := \tau(H')$ and $\tau_1(H) := \zeta_{H'} + \tau(H'')$, that are the only pair of times realizing the diameter $D(H)$ under Π_r, as already shown. Observe that under Π_r,

$$\mathcal{M}_{\tau_0(H),\tau_1(H)}(H) = \sum_{a \in \mathcal{J}_{0,\tau'}} \delta_{(r-a, \Lambda(\overrightarrow{H}^a), \Lambda(\overleftarrow{H}^a))} + \mathcal{M}_{0,\tau(H'')}(H'') \, .$$

where we recall here that Λ reverses excursions at their lifetime and that Λ is invariant under $\mathbf{N}_a$ and $\mathbf{P}^x_a$. Thus, basic results on Poisson point measures and an easy calculation shows that $\mathcal{M}_{\tau_0(H),\tau_1(H)}(H)$ is a Poisson point measure whose intensity is given by (3.35) in Theorem 3.2 (iv), which completes the proof of 3.2 (iv) because H under Π_r has law $\mathbf{Q}_r$ and thus $\mathcal{M}_{\tau_0(H),\tau_1(H)}(H)$ under Π_r has the same law as $\mathcal{M}_{\tau_0,\tau_1}$ under $\mathbf{Q}_r$. This completes the proof of Theorem 3.2. ■

3.3 Total height and diameter of normalized stable trees.

3.3.1 Preliminary results.

In this section, we gather general results that are used to prove Proposition 3.3. *Unless the contrary is explicitly mentioned, Ψ is a general branching mechanism of the form (3.13) that satisfies (3.14).* We first introduce the following function

$$\forall\, a, \lambda \in (0,\infty), \quad w_\lambda(a) := \mathbf{N}\big[1 - \mathbf{1}_{\{\Gamma \le a\}} e^{-\lambda\zeta}\big] \, . \tag{3.107}$$

For all fixed $\lambda \in (0,\infty)$, note that $a \mapsto w_\lambda(a)$ is non-increasing, that $\lim_{a\to 0} w_\lambda(a) = \infty$ and by (3.21) $\lim_{a\to\infty} w_\lambda(a) = \mathbf{N}[1 - e^{-\lambda\zeta}] = \Psi^{-1}(\lambda)$. As proved by Le Gall [81], Section II.3 (in the more general context of superprocesses) $w_\lambda(a)$ is the only solution of the following integral equation,

$$\forall\, a, \lambda \in (0,\infty), \quad \int_{w_\lambda(a)}^{\infty} \frac{du}{\Psi(u) - \lambda} = a \tag{3.108}$$

that makes sense thanks to (3.14).

Let us next consider H under $\mathbf{P}$ and recall from (3.16) that $\mathbf{P}^x$ stands for the law of $H_{\cdot \wedge T_x}$ where $T_x = \inf\{t \in \mathbb{R}_+ : X_t = -x\}$. Recall from (3.19) that $\sum_{i \in \mathcal{I}} \delta_{(-I_{a_i}, H^i)}$ stands for the decomposition of H into excursions above 0; thus, the excursions of $H_{\cdot \wedge T_x}$ above 0 are the H^i where $i \in \mathcal{I}$ is such that

$-I_{a_i} \in [0,x]$. Elementary results on Poisson point processes then imply the following:

$$\begin{aligned} \mathbf{E}_a^x\big[e^{-\lambda\zeta}\big] &= \mathbf{E}^x\big[e^{-\lambda\zeta}\mathbf{1}_{\{\Gamma\leq a\}}\big] \\ &= \mathbb{E}\Big[\exp\Big(-\sum_{i\in\mathcal{I}}\lambda\zeta_{H^i}\mathbf{1}_{[0,x]}(-I_{a_i})\Big)\mathbf{1}_{\big\{\Gamma(H^i)\leq a\,,\,i\in\mathcal{I}:-I_{a_i}\leq x\big\}}\Big] \\ &= \exp\big(-xw_\lambda(a)\big)\ . \end{aligned} \tag{3.109}$$

We first prove the following lemma.

Lemma 3.15. *Let Ψ be a branching mechanism of the form (3.13) that satisfies (3.14). Recall from (3.107) the definition of $w_\lambda(a)$. First observe that for all $a,\lambda\in(0,\infty)$,*

$$\partial_a w_\lambda(a) = \lambda - \Psi(w_\lambda(a)) \quad \textit{and} \quad \int_{w_\lambda(a)}^{\infty}\frac{du}{(\Psi(u)-\lambda)^2} = \frac{\partial_\lambda w_\lambda(a)}{\Psi(w_\lambda(a))-\lambda}\ . \tag{3.110}$$

Recall from (3.24) the definition of the function v. Then, for all $a,\lambda\in(0,\infty)$,

$$\lim_{\lambda\to 0+} w_\lambda(a) = v(a) \quad \textit{and} \quad v(a)\leq w_\lambda(a) = v(a)+\mathbf{N}_a\big[1-e^{-\lambda\zeta}\big]\leq v(a)+\Psi^{-1}(\lambda)\ , \tag{3.111}$$

where we recall from (3.34) the notation $\mathbf{N}_a$. Then, for all $r_1\geq r_0>0$, we get

$$\int_{r_0}^{r_1} da\,\Psi'(w_\lambda(a)) = \log\frac{\Psi(w_\lambda(r_0))-\lambda}{\Psi(w_\lambda(r_1))-\lambda} \quad \textit{and} \quad \int_{r_0}^{r_1} da\,\Psi'(v(a)) = \log\frac{\Psi(v(r_0))}{\Psi(v(r_1))}\ . \tag{3.112}$$

Proof. Note that (3.110) and (3.111) are easy consequences of resp. (3.108) and the definition (3.107). Let us first prove the first equality of (3.112): to that end we use the change of variable $u = w_\lambda(a)$, λ being fixed. Then, by (3.110), $-du/(\Psi(u)-\lambda) = da$, and we get

$$\int_{r_0}^{r_1} da\,\Psi'(w_\lambda(a)) = \int_{w_\lambda(r_1)}^{w_\lambda(r_0)} du\,\frac{\Psi'(u)}{\Psi(u)-\lambda} = \log\frac{\Psi(w_\lambda(r_0))-\lambda}{\Psi(w_\lambda(r_1))-\lambda}\ ,$$

which implies the second equality in (3.112) as $\lambda\to 0$ by (3.111). ■

Proposition 3.16. *Let Ψ be a branching mechanism of the form (3.13) that satisfies (3.14). Let $r\in(0,\infty)$. Recall from (3.30) the definition of $\mathbf{N}_r^\Gamma$ and recall from (3.107) the definition of $w_\lambda(a)$. Then for all $\lambda\in(0,\infty)$, we first get*

$$\mathbf{N}_r^\Gamma\big[e^{-\lambda\zeta}\big] = \exp\Big(-\int_0^r da\,\big(\Psi'(w_\lambda(a))-\Psi'(v(a))\big)\Big) = \frac{\Psi(w_\lambda(r))-\lambda}{\Psi(v(r))}\ . \tag{3.113}$$

Set $q_\lambda(y,r) := \mathbf{N}_r^\Gamma\big[e^{-\lambda\zeta}\mathbf{1}_{\{D>2y\}}\big]$. Then for all $y\in(\frac{1}{2}r,r)$, we have

$$q_\lambda(y,r) = \frac{\Psi(w_\lambda(r))-\lambda}{\Psi(v(r))}\Big(1-\frac{\big(\Psi(w_\lambda(y))-\lambda\big)^2}{\big(\Psi(w_\lambda(2y-r))-\lambda\big)\big(\Psi(w_\lambda(r))-\lambda\big)}\Big)\ . \tag{3.114}$$

If $y\leq\frac{1}{2}r$, then $q_\lambda(y,r) = \mathbf{N}_r^\Gamma\big[e^{-\lambda\zeta}\big]$ and if $y>r$, then $q_\lambda(y,r)=0$.

Proof. Recall from (3.84) the notation $\mathcal{M}_{0,\tau}$ and recall from (3.88) the notation $\Delta_{b,\overleftarrow{H},\overrightarrow{H}}$. Then, for all $r,y,\lambda\in(0,\infty)$, we get $\mathbf{N}_r^\Gamma$-a.s.

$$e^{-\lambda\zeta}\mathbf{1}_{\{D\leq 2y\}} = \exp\Big(-\lambda\sum_{a\in\mathcal{J}_{0,\tau}}\big(\zeta_{\overleftarrow{H}^a}+\zeta_{\overrightarrow{H}^a}\big)\Big)\mathbf{1}_{\big\{\forall a\in\mathcal{J}_{0,\tau}\,:\,\Delta_{a,\overleftarrow{H}^a,\overrightarrow{H}^a}\leq 2y\big\}}\ .$$

Lemma 3.11 asserts that under $\mathbf{N}_r^\Gamma$, $\mathcal{M}_{0,\tau}$ is a Poisson point measure with intensity $\mathbf{n}_r$ given by (3.85). Thus, elementary results on Poisson point measures imply that

$$\mathbf{N}_r^\Gamma\big[e^{-\lambda\zeta}\mathbf{1}_{\{D\le 2y\}}\big] = \exp\Big(-\underbrace{\int \mathbf{n}_r(da d\overleftarrow{H} d\overrightarrow{H})\big(1-\mathbf{1}_{\{\Delta_{a,\overleftarrow{H},\overrightarrow{H}}\le 2y\}}e^{-\lambda\zeta_{\overleftarrow{H}}-\lambda\zeta_{\overrightarrow{H}}}\big)}_{K}\Big).$$

Recall that the total mass of $\mathbf{P}_a^x$ is $e^{-xv(a)}$ and recall (3.109). Thus,

$$K = \int_0^r da\, 2\beta\, \mathbf{N}_a\big[1-\mathbf{1}_{\{\Gamma\le 2y-a\}}e^{-\lambda\zeta}\big] + \int_0^r da\int_{(0,\infty)}\pi(dz)\, z\big(e^{-zv(a)} - e^{-zw_\lambda(a\wedge(2y-a))}\big).$$

Now observe that

$$\mathbf{N}_a\big[1-\mathbf{1}_{\{\Gamma\le 2y-a\}}e^{-\lambda\zeta}\big] = \mathbf{N}\big[1-\mathbf{1}_{\{\Gamma\le a\wedge(2y-a)\}}e^{-\lambda\zeta}\big] - \mathbf{N}[\mathbf{1}_{\{\Gamma>a\}}] = w_\lambda\big(a\wedge(2y-a)\big)-v(a)\ .$$

Consequently,

$$\mathbf{N}_r^\Gamma\big[e^{-\lambda\zeta}\mathbf{1}_{\{D\le 2y\}}\big] = \exp\Big(-\int_0^r da\,\big(\Psi'(w_\lambda(a\wedge(2y-a)))-\Psi'(v(a))\big)\Big). \tag{3.115}$$

Then observe that if $y>r$, the $\mathbf{N}_r^\Gamma\big[e^{-\lambda\zeta}\mathbf{1}_{\{D\le 2y\}}\big] = \mathbf{N}_r^\Gamma\big[e^{-\lambda\zeta}\big]$ because $D\le 2\Gamma$. This combined with (3.115) entails the first equality of (3.113). Then, use (3.112) in Lemma 3.15 to get for any $\varepsilon\in(0,r)$,

$$\int_\varepsilon^r da\,\big(\Psi'(w_\lambda(a))-\Psi'(v(a))\big) = \log\frac{\Psi(v(r))}{\Psi(w_\lambda(r))-\lambda} - \log\frac{\Psi(v(\varepsilon))}{\Psi(w_\lambda(\varepsilon))-\lambda}.$$

This show that $\varepsilon\mapsto\Psi(v(\varepsilon))/(\Psi(w_\lambda(\varepsilon))-\lambda)$ is increasing and tends to a finite constant $C_\lambda\in(0,\infty)$ as $\varepsilon\to 0$. Thus, $C_\lambda^{-1}\Psi(v(r))\mathbf{N}_r^\Gamma\big[e^{-\lambda\zeta}\big]=\Psi(w_\lambda(r))-\lambda$, which is equal to $-\partial_r w_\lambda(r)$ by (3.110) in Lemma 3.15. Then recall from (3.24) that $\mathbf{N}(\Gamma\in dr)=\Psi(v(r))\,dr$; thus by (3.26) and the fact that $w_\lambda(r)$ tends to $\Psi^{-1}(\lambda)$ as $r\to\infty$, we get for all $b\in(0,\infty)$,

$$\begin{aligned} w_\lambda(b)-\Psi^{-1}(\lambda) &= \int_b^\infty dr\, C_\lambda^{-1}\Psi(v(r))\mathbf{N}_r^\Gamma\big[e^{-\lambda\zeta}\big] = C_\lambda^{-1}\mathbf{N}\big[e^{-\lambda\zeta}\mathbf{1}_{\{\Gamma>b\}}\big] \\ &= C_\lambda^{-1}\big(\mathbf{N}\big[1-\mathbf{1}_{\{\Gamma\le b\}}e^{-\lambda\zeta}\big]-\mathbf{N}\big[1-e^{-\lambda\zeta}\big]\big) = C_\lambda^{-1}\big(w_\lambda(b)-\Psi^{-1}(\lambda)\big). \end{aligned}$$

This implies that $C_\lambda=1$, which completes the proof of (3.113).

We next assume that $y\in(\frac{1}{2}r,r)$. Observe that $a\wedge(2y-a)=a$ if $a\in(0,y)$ and that $a\wedge(2y-a)=2y-a$ if $a\in(y,r)$. By (3.115) and (3.113), we then get

$$q_\lambda(y,r) = \mathbf{N}_r^\Gamma\big[e^{-\lambda\zeta}\big] - \mathbf{N}_r^\Gamma\big[e^{-\lambda\zeta}\mathbf{1}_{\{D\le 2y\}}\big] = \frac{\Psi(w_\lambda(r))-\lambda}{\Psi(v(r))}\Big(1-e^{-\int_y^r da\,\left(\Psi'(w_\lambda(2y-a))-\Psi'(w_\lambda(a)\right)}\Big),$$

which easily implies (3.114) by (3.112) in Lemma 3.15 since

$$\int_y^r da\Psi'(w_\lambda(2y-a)) = \int_{2y-r}^y da\Psi'(w_\lambda(a)) = \log\frac{\Psi(w_\lambda(2y-r))-\lambda}{\Psi(w_\lambda(y))-\lambda} \text{ and } \int_y^r da\Psi'(w_\lambda(a)) = \log\frac{\Psi(w_\lambda(y))-\lambda}{\Psi(w_\lambda(r))-\lambda}.$$

The other statements of the lemma follow immediately. ■

Proposition 3.17. *Let Ψ be a branching mechanism of the form (3.13) that satisfies (3.14). For all $y,z,\lambda\in(0,\infty)$, we have*

$$\begin{aligned} \mathrm{L}_\lambda(y,z) &:= \mathbf{N}\big[e^{-\lambda\zeta}\mathbf{1}_{\{D>2y\,;\,\Gamma>z\}}\big] \\ &= w_\lambda(y\vee z) - \Psi^{-1}(\lambda) - \mathbf{1}_{\{z\le 2y\}}\big(\Psi(w_\lambda(y))-\lambda\big)^2\int_{w_\lambda(y\wedge(2y-z))}^\infty \frac{du}{\big(\Psi(u)-\lambda\big)^2} \\ &= w_\lambda(y\vee z) - \Psi^{-1}(\lambda) - \mathbf{1}_{\{z\le 2y\}}\big(\Psi(w_\lambda(y))-\lambda\big)^2\frac{\partial_\lambda w_\lambda\big(y\wedge(2y-z)\big)}{\Psi\big(w_\lambda(y\wedge(2y-z))\big)-\lambda}\ . \end{aligned} \tag{3.116}$$

Proof. Recall notation $q_\lambda(y,r)$ from Proposition 3.16, which asserts that $q_\lambda(y,r)=0$ if $r<y$ and that $\Psi(v(r))q_\lambda(y,r)=-\partial_r w_\lambda(r)$, if $r\geq 2y$. Then, by (3.26), we get

$$\mathrm{L}_\lambda(y,z)=\int_z^\infty dr\,\Psi(v(r))q_\lambda(y,r)=\mathbf{1}_{\{z\leq 2y\}}\int_{z\vee y}^{2y} dr\,\Psi(v(r))q_\lambda(y,r)+\int_{z\vee 2y}^{\infty} dr\,\Psi(v(r))q_\lambda(y,r). \quad (3.117)$$

Since for all $r>z\vee 2y$, $\Psi(v(r))q_\lambda(y,r)=-\partial_r w_\lambda(r)$ and since $\lim_{r\to\infty} w_\lambda(r)=\Psi^{-1}(\lambda)$, we get

$$\int_{z\vee 2y}^{\infty} dr\,\Psi(v(r))q_\lambda(y,r)=w_\lambda(z\vee 2y)-\Psi^{-1}(\lambda)\,. \quad (3.118)$$

We next assume that $z\in(y,2y)$. By (3.114) and since $\Psi(w_\lambda(r))-\lambda=-\partial_r w_\lambda(r)$, we get

$$\begin{aligned}\int_z^{2y} dr\,\Psi(v(r))q_\lambda(y,r) &= -\int_z^{2y} dr\,\partial_r w_\lambda(r)-\big(\Psi(w_\lambda(y))-\lambda\big)^2\int_z^{2y}\frac{dr}{\Psi(w_\lambda(2y-r))-\lambda}\\ &= w_\lambda(z)-w_\lambda(2y)-\big(\Psi(w_\lambda(y))-\lambda\big)^2\int_0^{2y-z}\frac{dr}{\Psi(w_\lambda(r))-\lambda}\\ &= w_\lambda(z)-w_\lambda(2y)-\big(\Psi(w_\lambda(y))-\lambda\big)^2\int_{w_\lambda(2y-z)}^{\infty}\frac{du}{\big(\Psi(u)-\lambda\big)^2}\,,\end{aligned}$$

with the change of variable $u=w_\lambda(r)$ in the last line. This combined with (3.118) easily entails the first equality in (3.116). The second one follows from (3.110) in Lemma 3.15. ■

3.3.2 Proof of Proposition 3.3.

In this section, we fix $\gamma\in(1,2]$ and we take

$$\Psi(\lambda)=\lambda^\gamma,\quad \lambda\in\mathbb{R}_+\,.$$

Recall from (3.107) the definition of $w_\lambda(a)$. We then set

$$\forall y\in(0,\infty),\quad w(y):=w_1(y)\,. \quad (3.119)$$

Note that w satisfies (3.45) that is (3.108) with $\lambda=1$. By an easy change of variable (3.108) implies that

$$\forall a,\lambda\in(0,\infty),\quad w_\lambda(a)=\lambda^{\frac{1}{\gamma}}w\big(a\lambda^{\frac{\gamma-1}{\gamma}}\big)\,. \quad (3.120)$$

Recall from Proposition 3.17 the definition of $\mathrm{L}_\lambda(y,z)$. Then observe that the scaling property (3.44) entails (3.46). Moreover (3.47) follows from a simple change of variable. Next note from (3.120) that

$$\partial_\lambda w_\lambda(a)=\tfrac{1}{\gamma}\lambda^{\frac{1}{\gamma}-1}w\big(a\lambda^{\frac{\gamma-1}{\gamma}}\big)+\tfrac{\gamma-1}{\gamma}aw'\big(a\lambda^{\frac{\gamma-1}{\gamma}}\big).$$

This, combined with the fact that $-w'(y)=-\partial_y w_1(y)=w(y)^\gamma-1$, implies

$$\frac{\partial_\lambda w_1(y)}{w(y)^\gamma-1}=\frac{1}{\gamma}\frac{w(y)}{w(y)^\gamma-1}-\frac{\gamma-1}{\gamma}y\,,$$

which implies (3.48) thanks to the second equality in (3.116) in Proposition 3.17. This completes the proof of Proposition 3.3.

3.4 Proof of Theorems 3.5 and 3.7.

3.4.1 Preliminary results.

In this section we prove several estimates that are used in the proof of Theorems 3.5 and 3.7. We fix $\gamma\in(1,2]$ and we take $\Psi(\lambda)=\lambda^\gamma$, $\lambda\in\mathbb{R}_+$.

Laplace transform. We next introduce the following notation for Laplace transforms on $\mathbb{R}_+$ of Lebesgue integrable functions: for all measurable functions $f:\mathbb{R}_+\to\mathbb{R}$ such that there exists $\lambda_0\in\mathbb{R}_+$ with

$$\int_0^\infty dx\, e^{-\lambda_0 x}|f(x)| < \infty, \quad \text{we then set} \quad \mathcal{L}_\lambda(f) := \int_0^\infty dx\, e^{-\lambda x} f(x), \quad \lambda\in[\lambda_0,\infty) ,$$

which is well-defined. The function $\lambda\in[\lambda_0,\infty)\mapsto \mathcal{L}_\lambda(f)$ is the *Laplace transform of f*. We shall need the following lemma.

Lemma 3.18. *Let $f, g_n, h_n : \mathbb{R}_+ \to \mathbb{R}_+$, $n\in\mathbb{N}$, be continuous and nonnegative functions. We set $f_n := g_n - h_n$. Let $(q_n)_{n\geq 0}$ be a real valued sequence. We make the following assumptions.*

$$\exists \lambda_0\in\mathbb{R}_+ : \quad \int_0^\infty dx\, e^{-\lambda_0 x} f(x) < \infty \quad \textit{and} \quad \sum_{n\geq 0} |q_n| \int_0^\infty dx\, e^{-\lambda_0 x}\big(g_n(x)+h_n(x)\big) < \infty . \qquad (a)$$

This makes sense of the sum $\sum_{n\geq 0} q_n \mathcal{L}_\lambda(f_n)$ for all $\lambda\in[\lambda_0,\infty)$ and we assume that

$$\forall \lambda\in[\lambda_0,\infty), \quad \mathcal{L}_\lambda(f) = \sum_{n\geq 0} q_n \mathcal{L}_\lambda(f_n) . \qquad (b)$$

We furthermore assume

$$\forall x\in\mathbb{R}_+, \quad \sum_{n\geq 0} |q_n| \Big(\sup_{y\in[0,x]} g_n(y) + \sup_{y\in[0,x]} h_n(y)\Big) < \infty . \qquad (c)$$

Then,

$$\forall x\in\mathbb{R}_+, \quad f(x) = \sum_{n\geq 0} q_n f_n(x) ,$$

where the sum in the right member makes sense thanks to (c)*.*

Proof. We denote by $(\cdot)^+$ and $(\cdot)^-$ resp. the positive and negative part functions. Assumption (c) ensures that the following functions are well-defined for all $x\in\mathbb{R}_+$, continuous on $\mathbb{R}_+$ and nonnegative:

$$G := f + \sum_{n\geq 0} (q_n)^- g_n + (q_n)^+ h_n \quad \text{and} \quad H := \sum_{n\geq 0} (q_n)^+ g_n + (q_n)^- h_n .$$

Since the functions are nonnegative, for all $\lambda\in[\lambda_0,\infty)$, we get

$$\mathcal{L}_\lambda(G) = \mathcal{L}_\lambda(f) + \sum_{n\geq 0} (q_n)^- \mathcal{L}_\lambda(g_n) + (q_n)^+ \mathcal{L}_\lambda(h_n) \quad \text{and} \quad \mathcal{L}_\lambda(H) = \sum_{n\geq 0} (q_n)^+ \mathcal{L}_\lambda(g_n) + (q_n)^- \mathcal{L}_\lambda(h_n).$$

By Assumption (a), $\mathcal{L}_\lambda(G)$ and $\mathcal{L}_\lambda(H)$ are finite quantities for all $\lambda\geq\lambda_0$. Assumption (b) then entails that $\mathcal{L}_\lambda(G) = \mathcal{L}_\lambda(H)$, for all for all $\lambda\geq\lambda_0$: this implies that the Laplace transform of the finite Borel measures $e^{-\lambda_0 x}G(x)dx$ and $e^{-\lambda_0 x}H(x)dx$ are equal. Consequently, these measures are equal. Thus $G = H$ Lebesgue-almost everywhere. Since G and H are continuous, $G = H$ everywhere, which implies the desired result. ■

Estimates for stable distributions. Let $(\Omega,\mathcal{F},\mathbf{P})$ be an auxiliary space. Let $S:\Omega\to\mathbb{R}_+$ be a spectrally positive $\frac{\gamma-1}{\gamma}$-stable random variable such that

$$\forall \lambda\in\mathbb{R}+, \quad \mathbf{E}\big[e^{-\lambda S}\big] = \int_0^\infty dx\, s_\gamma(x)\exp(-\lambda x) = \exp\big(-\gamma\lambda^{\frac{\gamma-1}{\gamma}}\big), \qquad (3.121)$$

where we recall from (3.52) that $s_\gamma : \mathbb{R}_+ \to \mathbb{R}_+$ is the continuous version of the density of the $\frac{\gamma-1}{\gamma}$-stable distribution. We recall here from Ibragimov & Chernin [69] (see also Chambers, Mallows & Stuck [43] formula (2.1) p. 341 or Zolotarev [103]) the following representation of such a $\frac{\gamma-1}{\gamma}$-stable law: to that end, we first set

$$\forall v \in (-\pi, \pi), \quad m_\gamma(v) = \left(\frac{\gamma \sin\left(\frac{\gamma-1}{\gamma} v\right)}{\sin v} \right)^{\gamma-1} \frac{\gamma \sin\left(\frac{1}{\gamma} v\right)}{\sin v}. \tag{3.122}$$

Let V, W be two independent random variables defined on $(\Omega, \mathscr{F}, \mathbf{P})$ such that V is uniformly distributed on $[0, \pi]$ and such that W is exponentially distributed with mean 1. Then,

$$S \overset{\text{(law)}}{=} \left(\frac{m_\gamma(V)}{W} \right)^{\frac{1}{\gamma-1}},$$

which easily implies that

$$\forall x \in (0, \infty), \quad s_\gamma(x) = \frac{\gamma-1}{\pi} x^{-\gamma} \int_0^\pi dv \, m_\gamma(v) \exp\left(-x^{-(\gamma-1)} m_\gamma(v)\right). \tag{3.123}$$

We have $m_\gamma(-v) = m_\gamma(v)$ and $m_\gamma(0) = (\gamma-1)^{\gamma-1}$. Moreover, the function m_γ is increasing on $[0, \pi)$ and $m_\gamma(v)/m_\gamma(0) = 1 + \frac{\gamma-1}{2\gamma} v^2 + \mathcal{O}_\gamma(v^4)$.

As proved in Theorem 2.5.2 in Zolotarev [103], by an extension of Laplace's method (proved in Zolotarev [103], Lemma 2.5.1, p. 97) (3.121) yields the asymptotic expansion (3.53) that can be rewritten as follows: recall from (3.53) the definition of the sequence $(S_n)_{n\geq 1}$; then set

$$\forall x \in (0, \infty) \quad b(x) = \left(\frac{\gamma-1}{x} \right)^{\gamma-1} \quad \text{and} \quad S_n^* := \left(2\pi\left(1-\tfrac{1}{\gamma}\right)\right)^{-\frac{1}{2}} (\gamma-1)^{\frac{\gamma+1}{2} - n(\gamma-1)} S_n, \quad n \geq 0, \tag{3.124}$$

where recall that $S_0 = 1$. Then, for all positive integers N, as $x \to 0$, we have

$$s_\gamma(x) = \sum_{0 \leq n < N} S_n^* \, x^{n(\gamma-1) - \frac{\gamma+1}{2}} \, e^{-b(x)} + \mathcal{O}_{N,\gamma}\left(x^{N(\gamma-1) - \frac{\gamma+1}{2}} e^{-b(x)} \right). \tag{3.125}$$

For all $a \in \mathbb{R}$, we next set

$$\forall x \in \mathbb{R}_+ \quad J_a(x) := \int_0^x dy \, y^a e^{-b(y)} \tag{3.126}$$

An integration by parts entails

$$\forall a \in \mathbb{R} \backslash \{-\gamma\}, \; \forall x \in \mathbb{R}_+, \quad J_a(x) = (\gamma-1)^{-\gamma} x^{a+\gamma} e^{-b(x)} - (\gamma-1)^{-\gamma} (a+\gamma) J_{a+\gamma-1}(x), \tag{3.127}$$

which proves that $J_a(x) = \mathcal{O}_\gamma(x^{a+\gamma} e^{-b(x)})$ as $x \to 0$. This also entails the following lemma.

Lemma 3.19. *Let $\gamma \in (1, 2]$. Let $a \in \mathbb{R}$. We assume that $-(a+1)/(\gamma-1)$ is not a positive integer. Recall from (3.124) the definition of the function b and from (3.126) the definition of the function J_a. Then, we set*

$$\forall q \in \mathbb{N} \backslash \{0\}, \quad c_q(a, \gamma) := (-1)^q (\gamma-1)^{-(q+1)\gamma} \prod_{1 \leq k \leq q} \left(a+1+k(\gamma-1) \right), \tag{3.128}$$

with the convention that $c_0(a, \gamma) = (\gamma-1)^{-\gamma}$. Then, for all positive integers p,

$$J_a(x) = \sum_{0 \leq q < p} c_q(a, \gamma) \, x^{a+\gamma+q(\gamma-1)} \, e^{-b(x)} + (\gamma-1)^\gamma c_p(a, \gamma) \, J_{a+p(\gamma-1)}(x). \tag{3.129}$$

This implies that for all positive integers p, as $x \to 0$,

$$x^{-a-\gamma} e^{b(x)} J_a(x) = \sum_{0 \leq q < p} c_q(a, \gamma) \, x^{q(\gamma-1)} + \mathcal{O}_{p,a,\gamma}\left(x^{p(\gamma-1)} \right), \tag{3.130}$$

where $\mathcal{O}_{p,a,\gamma}$ depends on p, a and γ.

Proof: (3.129) follows from (3.127), by induction. Since $J_{a+p(\gamma-1)}(x) = \mathcal{O}_\gamma\big(x^{a+p(\gamma-1)+\gamma}e^{-b(x)}\big)$, (3.130) is an immediate consequence of (3.129). ■

We next prove the following lemma.

Lemma 3.20. *Let $\gamma\in(1,2]$. Recall from (3.52) (or from (3.121)) the definition of the density s_γ. Recall from (3.122) the definition of m_γ. We set for all $x\in\mathbb{R}_+$,*

$$\sigma^+(x) := \tfrac{(\gamma-1)^2}{\pi}x^{-2\gamma}\int_0^\pi dv\, m_\gamma(v)^2 e^{-x^{-(\gamma-1)}m_\gamma(v)} \tag{3.131}$$

$$\textit{and} \qquad \sigma^-(x) := \gamma x^{-1}s_\gamma(x) = \tfrac{\gamma(\gamma-1)}{\pi}x^{-\gamma-1}\int_0^\pi dv\, m_\gamma(v)e^{-x^{-(\gamma-1)}m_\gamma(v)} \; .$$

Then, the following holds true.

(i) σ^+ and σ^- are well-defined on $\mathbb{R}_+$, the function s_γ is differentiable on $\mathbb{R}_+$ and $s'_\gamma = \sigma^+ - \sigma^-$. Moreover, σ^+, σ^+ are continuous, nonnegative, Lebesgue integrable and for all $\lambda\in\mathbb{R}_+$,

$$\mathcal{L}_\lambda(\sigma^+) = \lambda e^{-\gamma\lambda^{\frac{\gamma-1}{\gamma}}} + \gamma\int_\lambda^\infty d\mu\, e^{-\gamma\mu^{\frac{\gamma-1}{\gamma}}} \quad \textit{and} \quad \mathcal{L}_\lambda(\sigma^-) = \gamma\int_\lambda^\infty d\mu\, e^{-\gamma\mu^{\frac{\gamma-1}{\gamma}}} \, , \tag{3.132}$$

which implies

$$\int_0^\infty dx\, |s'_\gamma(x)| < \infty \qquad \textit{and} \qquad \mathcal{L}_\lambda(s'_\gamma) = \lambda e^{-\gamma\lambda^{\frac{\gamma-1}{\gamma}}}, \quad \lambda\in\mathbb{R}_+. \tag{3.133}$$

(ii) There exist $A, x_0\in(0,\infty)$ such that

$$\forall x\in[0,x_0], \qquad \sigma^+(x) \;\textit{and}\; \sigma^-(x) \le Ax^{-\frac{3\gamma+1}{2}}e^{-b(x)} \, , \tag{3.134}$$

where we recall from (3.124) that $b(x) = \big((\gamma-1)/x\big)^{\gamma-1}$.

(iii) We define the real valued sequence $(T^*_n)_{n\ge 0}$ by

$$T^*_0 := (\gamma-1)^\gamma S^*_0 \quad \textit{and} \quad \forall n\ge 1, \quad T^*_n := (\gamma-1)^\gamma S^*_n + \big(n(\gamma-1) - \tfrac{3\gamma-1}{2}\big)S^*_{n-1} \, . \tag{3.135}$$

Then, for all positive integer N, as $x\to 0$, we have

$$s'_\gamma(x) = \sum_{0\le n<N} T^*_n\, x^{n(\gamma-1)-\frac{3\gamma+1}{2}}\, e^{-b(x)} + \mathcal{O}_{N,\gamma}\big(x^{N(\gamma-1)-\frac{3\gamma+1}{2}}e^{-b(x)}\big) \, . \tag{3.136}$$

Proof. We easily deduce from (3.123), that s_γ is differentiable on $\mathbb{R}_+$ and that $s'_\gamma = \sigma^+ - \sigma^-$. Using Fubini-Tonnelli and the change of variable $y = x^{-(\gamma-1)}m_\gamma(v)$, for fixed v, we get

$$\int_0^\infty dx\, \sigma^+(x) = \int_0^\infty dx\, \sigma^-(x) = \tfrac{\gamma}{\pi}\Gamma_e\big(\tfrac{\gamma}{\gamma-1}\big)\int_0^\pi dv\, m_\gamma(v)^{-\frac{1}{\gamma-1}} < \infty \, ,$$

since $m_\gamma(v) \ge m_\gamma(0) > 0$ on $[0,\pi)$ and $\lim_{v\to\pi} m_\gamma(v) = \infty$; here, Γ_e stands for Euler's gamma function. Thus, $\int_0^\infty dx\, |s'_\gamma(x)| < \infty$ and $\lambda\in\mathbb{R}_+ \mapsto \mathcal{L}_\lambda(s'_\gamma)$ is well-defined. Moreover, by Fubini,

$$\mathcal{L}_\lambda(s'_\gamma) = \int_0^\infty dx\, s'_\gamma(x)\int_x^\infty dy\, \lambda e^{-\lambda y} = \lambda\int_0^\infty dy\, e^{-\lambda y}\int_0^y dx\, s'_\gamma(x) = \lambda\mathcal{L}_\lambda(s_\gamma) \, ,$$

which completes the proof of (3.133). Next, by Fubini-Tonnelli, we get

$$\int_0^\infty dx\, e^{-\lambda x}x^{-1}s_\gamma(x) = \int_0^\infty dx\, s_\gamma(x)\int_\lambda^\infty d\mu\, e^{-\mu x} = \int_\lambda^\infty d\mu\, e^{-\gamma\mu^{\frac{\gamma-1}{\gamma}}} \, . \tag{3.137}$$

which implies that $\mathcal{L}_\lambda(\sigma^-) = \gamma\int_\lambda^\infty d\mu\, e^{-\gamma\mu^{\frac{\gamma-1}{\gamma}}}$, since $\sigma^-(x) = \gamma x^{-1}s_\gamma(x)$. This, combined with (3.133) entails (3.132), which completes the proof of (i).

Laplace's method easily implies that there exists $c_+, c_- \in (0,\infty)$ such that

$$\sigma^+(x) \underset{x\to 0}{\sim} c_+ x^{-\frac{3\gamma+1}{2}} e^{-b(x)} \quad \text{and} \quad \sigma^-(x) \underset{x\to 0}{\sim} c_- x^{-\frac{\gamma+3}{2}} e^{-b(x)} ,$$

which easily entails (3.134) and which completes the proof of (ii).

More generally, the asymptotic expansion (3.53) of s_γ is derived from (3.123) by an extension of Laplace's method proved in Zolotarev [103], Lemma 2.5.1, p. 97. When this method is applied to σ^+ and σ^-, one shows that σ^+ and σ^- have an asymptotic expansion whose general term is $x^{n(\gamma-1)-\frac{3\gamma+1}{2}}e^{-b(x)}$. Thus, there exists a sequence $(T_n^*)_{n\geq 0}$ such that (3.136) holds true. It remains to prove (3.135). To that end, for any $n\in\mathbb{N}$, we set $a_n := n(\gamma-1) - \frac{3\gamma+1}{2}$. By Lemma 3.19 we then get

$$\begin{aligned} s_\gamma(x) &= \sum_{0\leq n<N} T_n^* J_{a_n}(x) + \mathcal{O}_{N,\gamma}\big(J_{a_N}(x)\big) \\ &= \sum_{0\leq n<N}\sum_{0\leq q<N-n} T_n^* c_q(a_n,\gamma) x^{a_n+\gamma+q(\gamma-1)}e^{-b(x)} + \mathcal{O}_{N,\gamma}\big(x^{a_N+\gamma}e^{-b(x)}\big) \\ &= \sum_{0\leq n\leq p<N} T_n^* c_{p-n}(a_n,\gamma) x^{p(\gamma-1)-\frac{\gamma+1}{2}}e^{-b(x)} + \mathcal{O}_{N,\gamma}\big(x^{N(\gamma-1)-\frac{\gamma+1}{2}}e^{-b(x)}\big) , \end{aligned}$$

which implies that $S_p^* = \sum_{0\leq n\leq p} T_n^* c_{p-n}(a_n,\gamma)$, for all $p\in\mathbb{N}$. Then by (3.128), observe that

$$\begin{aligned} S_p^* &= c_0(a_p,\gamma)T_p^* + \sum_{0\leq n\leq p-1} T_n^* c_{p-n}(a_n,\gamma) \\ &= (\gamma-1)^{-\gamma}T_p^* - (\gamma-1)^{-\gamma}\big(p(\gamma-1) - \tfrac{3\gamma-1}{2}\big)\sum_{0\leq n\leq p-1} T_n^* c_{p-1-n}(a_n,\gamma) , \end{aligned}$$

which implies (3.135). This completes the proof of the lemma. ■

Proof of Proposition 3.6. Lemma 3.20 easily entails Proposition 3.6: indeed (3.133) entails (3.63). We then set

$$\forall n\in\mathbb{N}, \quad T_n := (\gamma-1)^{n(\gamma-1)}T_n^*/T_0^* ,$$

and we easily check that (3.135) entails (3.64) and that (3.136) implies (3.65). ■

We next introduce another function used in the asymptotic expansion of the height and the diameter of normalized stable tree.

Lemma 3.21. *Let $\gamma\in(1,2]$. Recall from (3.52) (or from (3.121)) the definition of s_γ. We then introduce the following functions: for all $x\in\mathbb{R}_+$,*

$$h^+(x) = (\gamma-1)\, x^{-1}s_\gamma(x), \quad h^-(x) = \tfrac{\gamma-1}{\gamma}x^{-1-\frac{1}{\gamma}}\int_0^x dy\, y^{\frac{1}{\gamma}-1}s_\gamma(y) \quad \textit{and} \quad \theta(x) = h^+(x) - h^-(x). \tag{3.138}$$

Then, the following holds true.

(i) h^+, h^- and θ are well-defined and continuous, h^+ and h^- are nonnegative and Lebesgue integrable, and for all $\lambda\in\mathbb{R}_+$, we have

$$\mathcal{L}_\lambda(h^+) = (\gamma-1)\int_\lambda^\infty d\mu\, e^{-\gamma\mu^{\frac{\gamma-1}{\gamma}}} \quad \textit{and} \quad \mathcal{L}_\lambda(h^-) = \mathcal{L}_\lambda(h^+) - \lambda^{\frac{1}{\gamma}}e^{-\gamma\lambda^{\frac{\gamma-1}{\gamma}}}, \tag{3.139}$$

which implies

$$\int_0^\infty dx\, |\theta(x)| < \infty \quad \textit{and} \quad \mathcal{L}_\lambda(\theta) = \lambda^{\frac{1}{\gamma}}e^{-\gamma\lambda^{\frac{\gamma-1}{\gamma}}}, \quad \lambda\in\mathbb{R}_+. \tag{3.140}$$

(ii) There exist $A, x_0 \in (0,\infty)$ such that

$$\forall x \in [0, x_0], \qquad h^+(x) \quad \textit{and} \quad h^-(x) \ \leq \ A x^{-\frac{\gamma+3}{2}} e^{-b(x)} \, , \tag{3.141}$$

where we recall from (3.124) that $b(x) = \big((\gamma-1)/x\big)^{\gamma-1}$.

(iii) Let $(V_n^*)_{n\geq 0}$ be a sequence of real numbers recursively defined by $V_0^* = (\gamma-1) S_0^*$ and for all $n \in \mathbb{N}$,

$$(\gamma-1)^{\gamma-1} V_{n+1}^* = (\gamma-1)^{\gamma} S_{n+1}^* + (\gamma-1)\big(n - \tfrac{1}{2} - \tfrac{1}{\gamma-1}\big) S_n^* - \big(n - \tfrac{1}{2} - \tfrac{1}{\gamma}\big) V_n^* \, . \tag{3.142}$$

Then for all positive integers N, as $x \to 0$, we get

$$\theta(x) = \sum_{0 \leq n < N} V_n^* \, x^{n(\gamma-1) - \frac{\gamma+3}{2}} e^{-b(x)} + \mathcal{O}_{N,\gamma}\big(x^{N(\gamma-1) - \frac{\gamma+3}{2}} e^{-b(x)}\big). \tag{3.143}$$

Proof. The fact that h^+ and h^- are well-defined is an easy consequence of the asymptotic expansion (3.125) of s_γ and observe that h^+, h^- can be continuously extended by the value 0 at $x = 0$. Let $\lambda \in \mathbb{R}_+$; by (3.137) we get $\mathcal{L}_\lambda(h^+) = (\gamma-1) \int_\lambda^\infty d\mu \, \exp(-\gamma \mu^{\frac{\gamma-1}{\gamma}})$. Thus when $\lambda = 0$, we get

$$\int_0^\infty dx \, h^+(x) = \mathcal{L}_0(h^+) = (\gamma-1) \int_0^\infty d\mu \, e^{-\gamma \mu^{\frac{\gamma-1}{\gamma}}} = \gamma^{-\frac{1}{\gamma-1}} \Gamma_e\big(\tfrac{\gamma}{\gamma-1}\big),$$

by an easy change of variable; here Γ_e stands for Euler's Gamma function. By Fubini-Tonnelli and several linear changes of variable, we get

$$\begin{aligned}
\mathcal{L}_\lambda(h^-) &= \tfrac{\gamma-1}{\gamma} \int_0^\infty dy \, y^{\frac{1}{\gamma}-1} s_\gamma(y) \int_y^\infty dx \, x^{-1-\frac{1}{\gamma}} e^{-\lambda x} = \tfrac{\gamma-1}{\gamma} \lambda^{\frac{1}{\gamma}} \int_0^\infty dy \, y^{\frac{1}{\gamma}-1} s_\gamma(y) \int_{\lambda y}^\infty d\mu \, \mu^{-1-\frac{1}{\gamma}} e^{-\mu} \\
&= \tfrac{\gamma-1}{\gamma} \lambda^{\frac{1}{\gamma}} \int_0^\infty dy \, y^{-1} s_\gamma(y) \int_\lambda^\infty d\nu \, \nu^{-1-\frac{1}{\gamma}} e^{-\nu y} = \tfrac{\gamma-1}{\gamma} \lambda^{\frac{1}{\gamma}} \int_\lambda^\infty d\nu \, \nu^{-1-\frac{1}{\gamma}} \int_0^\infty dy \, y^{-1} s_\gamma(y) e^{-\nu y} \\
&= \tfrac{\gamma-1}{\gamma} \lambda^{\frac{1}{\gamma}} \int_\lambda^\infty d\nu \, \nu^{-1-\frac{1}{\gamma}} \int_\nu^\infty d\mu \, e^{-\gamma \mu^{\frac{\gamma-1}{\gamma}}} = (\gamma-1) \lambda^{\frac{1}{\gamma}} \int_\lambda^\infty d\mu \, e^{-\gamma \mu^{\frac{\gamma-1}{\gamma}}} \big(\lambda^{-\frac{1}{\gamma}} - \mu^{-\frac{1}{\gamma}}\big) \\
&= (\gamma-1) \int_\lambda^\infty d\mu \, e^{-\gamma \mu^{\frac{\gamma-1}{\gamma}}} - (\gamma-1) \lambda^{\frac{1}{\gamma}} \int_\lambda^\infty d\mu \, \mu^{-\frac{1}{\gamma}} e^{-\gamma \mu^{\frac{\gamma-1}{\gamma}}} \\
&= (\gamma-1) \int_\lambda^\infty d\mu \, e^{-\gamma \mu^{\frac{\gamma-1}{\gamma}}} - \lambda^{\frac{1}{\gamma}} e^{-\gamma \lambda^{\frac{\gamma-1}{\gamma}}} .
\end{aligned}$$

Here we use (3.137) in the third line. When $\lambda = 0$, this proves that

$$\int_0^\infty dx \, h^-(x) = \gamma^{-\frac{1}{\gamma-1}} \Gamma_e\big(\tfrac{\gamma}{\gamma-1}\big) \, .$$

Thus, $\int_0^\infty dx \, |\theta(x)| < \infty$. Combined with (3.137), it also implies (3.140), which completes the proof of (i).

We then prove (ii) and (iii). To that end, we first observe that (3.125) implies that $x^{-1} s_\gamma(x) \sim S_0^* x^{-\frac{\gamma+3}{2}} e^{-b(x)}$ as $x \to 0$, which immediately entails (3.141) for h^+.

We next find the asymptotic expansion of h^- thanks to that of s_γ and thanks to Lemma 3.19. We first set $\alpha_n = \frac{1}{\gamma} - \frac{\gamma+3}{2} + n(\gamma-1)$. From (3.125) and Lemma 3.19, for all positive integer N, as $x \to 0$, we get

$$
\begin{aligned}
h^{-}(x) &= \sum_{0\le n<N} \tfrac{\gamma-1}{\gamma} S_n^* \, x^{-1-\frac{1}{\gamma}} J_{\alpha_n}(x) + \mathcal{O}_{N,\gamma}\big(x^{-1-\frac{1}{\gamma}} J_{\alpha_N}(x)\big) \\
&= \sum_{0\le n<N}\ \sum_{0\le q<N-n} \tfrac{\gamma-1}{\gamma} S_n^* c_q(\alpha_n,\gamma) x^{\alpha_n+\gamma-1-\frac{1}{\gamma}+q(\gamma-1)} e^{-b(x)} + \mathcal{O}_{N,\gamma}\big(x^{\alpha_N+\gamma-1-\frac{1}{\gamma}} e^{-b(x)}\big) \\
&= \sum_{0\le n<N}\ \sum_{0\le q<N-n} \tfrac{\gamma-1}{\gamma} S_n^* c_q(\alpha_n,\gamma) x^{(n+q+1)(\gamma-1)-\frac{\gamma+3}{2}} e^{-b(x)} + \mathcal{O}_{N,\gamma}\big(x^{(N+1)(\gamma-1)-\frac{\gamma+3}{2}} e^{-b(x)}\big) \\
&= \sum_{0\le p\le N} U_p \, x^{p(\gamma-1)-\frac{\gamma+3}{2}} e^{-b(x)} + \mathcal{O}_{N,\gamma}\big(x^{(N+1)(\gamma-1)-\frac{\gamma+3}{2}} e^{-b(x)}\big).
\end{aligned}
$$

where the sequence $(U_p)_{p\ge 0}$ is given by

$$
U_0 = 0, \quad \text{and} \quad U_p = \sum_{0\le n\le p-1} \tfrac{\gamma-1}{\gamma} S_n^* c_{p-1-n}(\alpha_n,\gamma), \quad p\ge 1.
$$

Observe that it implies (3.141) for h^-, which completes the proof of (ii). We next prove (iii): to that end observe that by (3.128), $c_{p-n}(\alpha_n,\gamma) = -(\gamma-1)^{-\gamma}\big(\frac{1}{\gamma} - \frac{\gamma+1}{2} + p(\gamma-1)\big) c_{p-1-n}(\alpha_n,\gamma)$. Thus we get

$$
\begin{aligned}
U_{p+1} &= \sum_{0\le n\le p} \tfrac{\gamma-1}{\gamma} S_n^* c_{p-n}(\alpha_n,\gamma) \ = \tfrac{\gamma-1}{\gamma} S_p^* c_0(\alpha_p,\gamma) \ + \sum_{0\le n\le p-1} \tfrac{\gamma-1}{\gamma} S_n^* c_{p-n}(\alpha_n,\gamma) \\
&= \tfrac{1}{\gamma}(\gamma-1)^{-(\gamma-1)} S_p^* - (\gamma-1)^{-\gamma}\big(\tfrac{1}{\gamma} - \tfrac{\gamma+1}{2} + p(\gamma-1)\big) \sum_{0\le n\le p-1} \tfrac{\gamma-1}{\gamma} S_n^* c_{p-1-n}(\alpha_n,\gamma) \\
&= \tfrac{1}{\gamma}(\gamma-1)^{-(\gamma-1)} S_p^* - (\gamma-1)^{-\gamma}\big(\tfrac{1}{\gamma} - \tfrac{\gamma+1}{2} + p(\gamma-1)\big) U_p \\
&= (\gamma-1)^{-(\gamma-1)}\Big(\tfrac{1}{\gamma} S_p^* - \big(p - \tfrac{1}{2} - \tfrac{1}{\gamma}\big) U_p\Big).
\end{aligned} \tag{3.144}
$$

We then set $V_p^* = (\gamma-1)S_p^* - U_p$ for all $p\in\mathbb{N}$, so that for all positive integer N, as $x\to 0$, (3.143) holds true. Moreover, by (3.144), easily entails that $(V_p^*)_{p\ge 0}$ satisfies (3.142), which completes the proof of the lemma. ■

Proof of Proposition 3.4. Lemma 3.21 easily entails Proposition 3.4. Indeed, (3.140) implies (3.55). We set

$$
\forall n\in\mathbb{N}, \quad V_n = (\gamma-1)^{n(\gamma-1)} V_n^*/V_0^* \, .
$$

Then, (3.142) entails (3.56) and (3.143) implies (3.57), which completes the proof of Proposition 3.4. ■

Lemma 3.22. *There exist $\lambda_0, A\in(0,\infty)$ such that*

$$
\forall \lambda\in[\lambda_0,\infty), \quad \int_\lambda^\infty d\mu \, e^{-\gamma\mu^{\frac{\gamma-1}{\gamma}}} \le A\lambda^{\frac{1}{\gamma}} e^{-\gamma\lambda^{\frac{\gamma-1}{\gamma}}} \, .
$$

Proof. Integration by part implies

$$
(\gamma-1)\int_\lambda^\infty d\mu \, e^{-\gamma\mu^{\frac{\gamma-1}{\gamma}}} = \lambda^{\frac{1}{\gamma}} e^{-\gamma\lambda^{\frac{\gamma-1}{\gamma}}} + \frac{1}{\gamma}\int_\lambda^\infty d\mu \, \mu^{-\frac{\gamma-1}{\gamma}} e^{-\gamma\mu^{\frac{\gamma-1}{\gamma}}} \le \lambda^{\frac{1}{\gamma}} e^{-\gamma\lambda^{\frac{\gamma-1}{\gamma}}} + \frac{1}{\gamma}\lambda^{-\frac{\gamma-1}{\gamma}}\int_\lambda^\infty d\mu \, e^{-\gamma\mu^{\frac{\gamma-1}{\gamma}}},
$$

which immediately entails the lemma. ■

Asymptotic expansion of $w-1$. Recall from (3.45) the definition of w. We next introduce

$$\forall y\in(0,\infty),\quad \phi(y):=w(y)-1,\quad \text{that satisfies}\quad \int_{\phi(y)}^{\infty}\frac{du}{(u+1)^{\gamma}-1}=y\ , \tag{3.145}$$

by (3.45). We easily see that $\lim_{y\to\infty}\phi(y)=0$ and $\lim_{y\to 0}\phi(y)=\infty$ and that ϕ is a C^{∞} decreasing function. The following lemma asserts that ϕ decreases exponentially fast as $y\to\infty$.

Lemma 3.23. *Let $\gamma\in(1,2]$. Let $\Psi(\lambda)=\lambda^{\gamma}$, $\lambda\in\mathbb{R}_+$. Recall from (3.145) the definition of ϕ. We set*

$$y_0:=\int_1^{\infty}\frac{du}{(u+1)^{\gamma}-1}\quad \textit{and}\quad \forall y\in[-1,\infty),\quad G(y):=\int_0^y\frac{du}{u}\,\frac{(u+1)^{\gamma}-1-\gamma u}{(u+1)^{\gamma}-1}\ . \tag{3.146}$$

Then,

$$\forall y\in[-1,1],\quad \exp(G(y))=1+\sum_{n\geq 1}A_ny^n\quad \textit{and}\quad 1+\sum_{n\geq 1}|A_n|<e^{\gamma-1}.$$

Moreover, for $y\in[y_0,\infty)$,

$$e^{\gamma y-C_0}\phi(y)=\exp\big(G(\phi(y))\big)=1+\sum_{n\geq 1}A_n\phi(y)^n\ , \tag{3.147}$$

where C_0 is given by (3.62).

Proof. We first introduce the inverse function of ϕ. Namely, for all $y\in(0,\infty)$, we set

$$\forall y\in(0,\infty),\quad F(y):=\int_y^{\infty}\frac{du}{(u+1)^{\gamma}-1}\ .$$

Observe that

$$F(y)=\int_1^{\infty}\frac{du}{(u+1)^{\gamma}-1}+\frac{1}{\gamma}\int_y^1\frac{du}{u}-\frac{1}{\gamma}\int_0^1\frac{du}{u}\,\frac{(u+1)^{\gamma}-1-\gamma u}{(u+1)^{\gamma}-1}+\frac{1}{\gamma}\int_0^y\frac{du}{u}\,\frac{(u+1)^{\gamma}-1-\gamma u}{(u+1)^{\gamma}-1},$$

which makes sense since $\frac{1}{u}\,\frac{(u+1)^{\gamma}-1-\gamma u}{(u+1)^{\gamma}-1}\to\frac{\gamma-1}{2}$ as $u\to 0+$. We then set

$$C_0:=\gamma\int_1^{\infty}\frac{du}{(u+1)^{\gamma}-1}-\int_0^1\frac{du}{u}\,\frac{(u+1)^{\gamma}-1-\gamma u}{(u+1)^{\gamma}-1}$$

and we get

$$\forall y\in(0,\infty),\quad \gamma F(y)=C_0-\log y+G(y),\quad \text{where}\quad G(y):=\int_0^y\frac{du}{u}\,\frac{(u+1)^{\gamma}-1-\gamma u}{(u+1)^{\gamma}-1}.$$

Since $F(\phi(y))=y$, this implies

$$\forall y\in(0,\infty),\quad \log\phi(y)=C_0-\gamma y+G(\phi(y))\ . \tag{3.148}$$

Let us show that $G(y)$ (and therefore $\exp(G(y))$) is analytic in a neighborhood of 0. We set

$$a_n=\frac{1}{\gamma}\binom{\gamma}{n+1}=\frac{(-1)^{n-1}}{(n+1)!}\prod_{k=1}^{n}|k-\gamma|=\frac{(\gamma-1)(-1)^{n-1}}{n(n+1)}\prod_{k=1}^{n-1}\Big(1-\frac{\gamma-1}{k}\Big),\quad n\geq 1.$$

We observe that $|a_n|<\frac{\gamma-1}{n(n+1)}$. Then for all $u\in[-1,1]$, we set

$$T(u):=\sum_{n\geq 1}|a_n|u^n\quad \text{and}\quad S(u):=\frac{(1+u)^{\gamma}-1-\gamma u}{\gamma u}=\sum_{n\geq 1}a_nu^n=-T(-u)$$

since $(-1)^{n-1}a_n = |a_n|$. The power series T and S are absolutely convergent for $|u| \leq 1$. Moreover, $|S(u)| \leq T(|u|) \leq T(1) = -S(-1) = \frac{\gamma-1}{\gamma} < 1$. Thus, for all $u \in [-1,1]$,

$$\frac{(1+u)^\gamma - 1 - \gamma u}{(1+u)^\gamma - 1} = \frac{S(u)}{1+S(u)} = \sum_{p\geq 1}(-1)^{p-1}S(u)^p = \sum_{n\geq 1}(-1)^{n-1} n B_n u^n$$

is analytic for $|u| \leq 1$, where

$$nB_n = \sum_{\substack{p_1,\ldots,p_n\geq 0\\ p_1+2p_2+\ldots+np_n=n}} \frac{(p_1+\ldots+p_n)!}{p_1!\ldots p_n!}|a_1|^{p_1}\ldots|a_n|^{p_n} \quad \geq 0\,.$$

Note that $\sum_{n\geq 1} nB_n = T(1)/(1-T(1)) = \gamma - 1 < 1$. Therefore,

$$\forall y \in [-1,1], \quad G(y) = \sum_{n\geq 1}(-1)^{n-1}B_n y^n\,,$$

is absolutely convergent and $|G(y)| \leq -G(-1) < \sum_{n\geq 1} nB_n = \gamma - 1 < 1$. Thus

$$\forall y \in [-1,1], \quad \exp(G(y)) = 1 + \sum_{n\geq 1} A_n y^n$$

where

$$A_n = \sum_{\substack{p_1,\ldots,p_n\geq 0\\ p_1+2p_2+\cdots+np_n=n}} \frac{(-B_1)^{p_1}\ldots(-B_n)^{p_n}}{p_1!\ldots p_n!}$$

and

$$1 + \sum_{n\geq 1}|A_n| \leq 1 + \sum_{n\geq 1}\sum_{\substack{p_1,\ldots,p_n\geq 0\\ p_1+2p_2+\cdots+np_n=n}} \frac{B_1^{p_1}\ldots B_n^{p_n}}{p_1!\ldots p_n!} = \exp(-G(-1)) < \exp(\gamma - 1)\,.$$

Observe that $\phi(y_0) = 1$. Then (3.147) follows from (3.148) for all $y \in [y_0, \infty)$. ■

We then derive from the previous lemma the following expansion of ϕ.

Lemma 3.24. *Let $\gamma \in (1,2]$. Let $\Psi(\lambda) = \lambda^\gamma$, $\lambda \in \mathbb{R}_+$. Recall from (3.145) the definition of ϕ; recall from (3.146) the definition of G and recall from (3.62) the definition of C_0. Then, we set*

$$\forall y \in [-1,1], \quad H(y) := \exp\big(C_0 + G(y)\big) \quad \textit{and} \quad \forall n \geq 1, \quad \beta_n := \frac{1}{n!}\frac{d^{n-1}}{dy^{n-1}}\big(H^n\big)\Big|_{y=0}\,. \tag{3.149}$$

Then, there exists $y_1 \in [y_0, \infty)$ such that

$$\sum_{n\geq 1}|\beta_n|e^{-\gamma n y_1} < \infty \quad \textit{and} \quad \forall y \in [y_1, \infty), \qquad \phi(y) = \sum_{n\geq 1}\beta_n e^{-\gamma n y}\,. \tag{3.150}$$

Here $\beta_1 = e^{C_0}$ and $\beta_2 = \frac{\gamma-1}{4}e^{C_0}$.

Proof. Lemma 3.23 shows that H has a power expansion whose radius of convergence is larger than 1. By Lagrange's inversion theorem (see for instance Whittaker & Watson [102], 7.32, pp. 132–133), there exists $x_0 \in (0,\infty)$ such that $\sum_{n\geq 1}|\beta_n|x_0^n < \infty$ and

$$\forall x \in [-x_0, x_0], \quad f(x) := \sum_{n\geq 1}\beta_n x^n \quad \text{satisfies} \quad f(x) = xH(f(x))\,.$$

For all $x, y \in \mathbb{R}_+$ in a neighborhood of 0, we next set $Q(x,y) = H(y)/(1-xH'(y))$. Observe that f is a solution of the differential equation $f'(x) = Q(x, f(x))$ in a neighborhood of 0. Then note that

$$\partial_x Q(0,0) = H(0)H'(0) = \tfrac{\gamma-1}{2} e^{2C_0} \quad \text{and} \quad \partial_y Q(0,0) = H'(0) = \tfrac{\gamma-1}{2} e^{C_0}.$$

Thus, there exists $x_0^* \in (0, x_0)$ such that Q is Lipschitz on $[-x_0^*, x_0^*]^2$. By the Cauchy-Lipschitz theorem, f is the unique solution on $[0, x_0^*]$ of the equation $y'(x) = Q(x, y(x))$ such that $y(0) = 0$.

We next recall from Lemma 3.23 that $\phi(y) = e^{-\gamma y} H(\phi(y))$, for all $y \in [y_0, \infty)$. For all $x \in (0, e^{-\gamma y_0}]$, we set $g(x) := \phi(-\frac{1}{\gamma}\log x)$ and $g(0) = 0$. Note that g is differentiable on $(0, e^{-\gamma y_0}]$ and that $g(x) = xH(g(x))$ for all $x \in [0, e^{-\gamma y_0}]$. Thus $g'(x) = Q(x, g(x))$, for all $(0, e^{-\gamma y_0}]$. This implies that $\lim_{0 \to 0+} x^{-1} g(x) = \lim_{0 \to 0+} g'(x) = H(0)$, which proves that g is a C^1 function satisfying the same differential equation as f on a neighborhood of 0, with the same initial value 0 at $x = 0$. Consequently, there exists $x_1 \in (0, x_0^* \wedge e^{-\gamma y_0}]$ such that $g(x) = f(x)$ for all $x \in [0, x_1]$ which implies (3.150) with $y_1 := -\frac{1}{\gamma}\log x_1$. The values of β_1 and β_2 are easily derived from (3.149). ■

We next derive from the previous lemma a similar asymptotic expansion for the function $\mathrm{L}_1(y, 0)$ that is connected to the diameter of γ-stable normalized trees.

Lemma 3.25. *Let $\gamma \in (1, 2]$. Let $\Psi(\lambda) = \lambda^\gamma$, $\lambda \in \mathbb{R}_+$. Recall from (3.49) the definition of $\mathrm{L}_1(y, 0)$ and recall from (3.62) the definition of C_0. Then, there exist $y_2 \in (0, \infty)$, and two real valued sequences $(\gamma_n)_{n \geq 2}$, $(\delta_n)_{n \geq 2}$ such that*

$$\gamma_2 = \tfrac{1}{2}\gamma(\gamma-1)e^{2C_0}, \quad \delta_2 = -\tfrac{1}{2}(\gamma+1)e^{2C_0} \quad \textit{and} \quad \sum_{n \geq 2} \big(n|\gamma_n| + |\delta_n|\big) e^{-\gamma n y_2} < \infty \tag{3.151}$$

and

$$\forall y \in [y_2, \infty), \quad \mathrm{L}_1(y, 0) = \sum_{n \geq 2} (n\gamma_n y + \delta_n) e^{-\gamma n y} \, . \tag{3.152}$$

Proof. Recall from (3.145) that $\phi(y) = w(y) - 1$, where w is defined by (3.45). Then, (3.49) implies the following:

$$\begin{aligned} L_1(y, 0) &= \phi(y) - \tfrac{1}{\gamma}\big[(1+\phi(y))^\gamma - 1\big](1+\phi(y)) + \tfrac{\gamma-1}{\gamma}\, y \,\big[(1+\phi(y))^\gamma - 1\big]^2 \\ &= \tfrac{\gamma-1}{\gamma}\, y \,\big[(1+\phi(y))^\gamma - 1\big]^2 - \tfrac{1}{\gamma}\big[(1+\phi(y))^{\gamma+1} - 1 - (\gamma+1)\phi(y)\big] \\ &= \gamma(\gamma-1) y \phi(y)^2 K(\phi(y)) - \tfrac{1}{2}(\gamma+1)\phi(y)^2 M(\phi(y)) \end{aligned} \tag{3.153}$$

where for all $u \in [-1, \infty)$ we have set

$$K(u) = \frac{\big((u+1)^\gamma - 1\big)^2}{(\gamma u)^2} \quad \text{and} \quad M(u) = \frac{(u+1)^{\gamma+1} - 1 - (\gamma+1)u}{\frac{1}{2}\gamma(\gamma+1)u^2}.$$

Recall from (3.149) the definition of H and recall from (3.147) that for all $y \in [y_0, \infty)$, $\phi(y) = e^{-\gamma y} H(\phi(y))$. This, combined with (3.153), entails that

$$\mathrm{L}_1(y, 0) = \gamma(\gamma-1) e^{-2\gamma y} H(\phi(y))^2 K(\phi(y)) - \tfrac{1}{2}(\gamma+1) e^{-2\gamma y} H(\phi(y))^2 M(\phi(y)) \, . \tag{3.154}$$

Recall from (3.149) the definition of the real valued sequence $(\beta_n)_{n \geq 1}$ that is such that $\sum_{n \geq 1} |\beta_n| x_1^n < \infty$, where $x_1 = e^{-\gamma y_1}$. We then set $f(x) = \sum_{n \geq 1} \beta_n x^n$, for all $x \in [-x_1, x_1]$. Next, recall from Lemma 3.23 that $e^{G(y)}$ has a power expansion on $[-1, 1]$; thus, so does H. Note that K and M have also a power expansion on $(-1, 1)$. Consequently there exists $x_2 \in (0, \infty)$ such for all $x \in [0, x_2]$,

$$\gamma(\gamma-1) H(f(x))^2 K(f(x)) = \sum_{n \geq 0} \gamma'_n x^n \quad \text{and} \quad -\tfrac{1}{2}(\gamma+1) H(f(x))^2 M(f(x)) = \sum_{n \geq 0} \delta'_n x^n, \tag{3.155}$$

with

$$\gamma'_0 = \gamma(\gamma-1)e^{2C_0}, \quad \delta'_0 = -\tfrac{1}{2}(\gamma+1)e^{2C_0} \quad \text{and} \quad \sum_{n\geq 0} \big(|\gamma'_n| + |\delta'_n|\big)x_2^n < \infty, \tag{3.156}$$

since $K(0) = M(0) = 1$ and since $H(0)^2 = e^{2C_0}$. Next by (3.150) in Lemma 3.24, we have $\phi(y) = f(e^{-\gamma y})$, for all $y \in [y_1, \infty)$. Then we set

$$y_2 := y_1 \wedge \big(-\tfrac{1}{\gamma}\log x_2\big) \quad \text{and} \quad \forall n \geq 2, \quad \gamma_n := n^{-1}\gamma'_{n-2}, \quad \delta_n := \delta'_{n-2}\,,$$

and (3.156) implies (3.151); (3.155) and (3.154) imply (3.152), which completes the proof of the lemma. ∎

3.4.2 Proof of Theorem 3.5.

We first set

$$\forall x \in (0,\infty), \quad f_\Gamma(x) := c_\gamma x^{-1-\frac{1}{\gamma}} \mathbf{N}_{\mathrm{nr}}\big(\Gamma > x^{-\frac{\gamma-1}{\gamma}}\big)\,. \tag{3.157}$$

Then, Proposition 3.3, (3.46), (3.47) and (3.49) imply for all $\lambda \in (0,\infty)$,

$$\mathcal{L}_\lambda(f_\Gamma) = \int_0^\infty dx\, e^{-\lambda x} f_\Gamma(x) = \mathrm{L}_\lambda(0,1) = \lambda^{\frac{1}{\gamma}} \mathrm{L}_1\big(0, \lambda^{\frac{\gamma-1}{\gamma}}\big) = \lambda^{\frac{1}{\gamma}}\big(w\big(\lambda^{\frac{\gamma-1}{\gamma}}\big) - 1\big) = \lambda^{\frac{1}{\gamma}}\phi\big(\lambda^{\frac{\gamma-1}{\gamma}}\big), \tag{3.158}$$

where recall from (3.145) that $\phi(y) = w(y) - 1$. We next use Lemma 3.24: let λ_1 be such that $\lambda_1^{\frac{\gamma-1}{\gamma}} = y_1$; then the sequence $(\beta_n)_{n\geq 0}$ satisfies

$$\forall \lambda \in [\lambda_1, \infty), \quad \sum_{n\geq 0} |\beta_n| \lambda^{\frac{1}{\gamma}} e^{-\gamma n \lambda^{\frac{\gamma-1}{\gamma}}} < \infty \quad \text{and} \quad \mathcal{L}_\lambda(f_\Gamma) = \sum_{n\geq 0} \beta_n \lambda^{\frac{1}{\gamma}} e^{-\gamma n \lambda^{\frac{\gamma-1}{\gamma}}}\,. \tag{3.159}$$

Recall from Lemma 3.21 the definition of the functions θ, h^+ and h^-. Then for all integer $n \geq 1$, and all $x \in \mathbb{R}_+$, we set

$$\theta_n(x) = n^{-\frac{\gamma+1}{\gamma-1}} \theta\big(n^{-\frac{\gamma}{\gamma-1}} x\big), \quad h_n^+(x) = n^{-\frac{\gamma+1}{\gamma-1}} h^+\big(n^{-\frac{\gamma}{\gamma-1}} x\big) \quad \text{and} \quad h_n^-(x) = n^{-\frac{\gamma+1}{\gamma-1}} h^-\big(n^{-\frac{\gamma}{\gamma-1}} x\big).$$

Lemma 3.21 implies that h_n^+, h_n^- are Lebesgue integrable, nonnegative and continuous. Moreover, $\theta_n = h_n^+ - h_n^-$. Consequently, θ_n is also nonnegative continuous and Lebesgue integrable, and (3.55) entails that $\mathcal{L}_\lambda(\theta_n) = \lambda^{\frac{1}{\gamma}} e^{-\gamma n \lambda^{\frac{\gamma-1}{\gamma}}}$. Thus, by (3.159)

$$\forall \lambda \in [\lambda_1, \infty), \quad \mathcal{L}_\lambda(f_\Gamma) = \sum_{n\geq 0} \beta_n \mathcal{L}_\lambda(\theta_n). \tag{3.160}$$

We next prove that the assumptions (a), (b), (c) of Lemma 3.18 hold true with

$$f := f_\Gamma, \quad f_n^+ := h_n^+, \quad f_n^- := h_n^-, \quad \text{and} \quad q_n := \beta_n\,.$$

To that end, we first observe that by an easy change of variable and by (3.139) in Lemma 3.21, we get

$$\forall \lambda \in (0,\infty),\ \forall n \geq 1, \qquad \mathcal{L}_\lambda(h_n^+) \ \text{ and } \ \mathcal{L}_\lambda(h_n^-) \ \leq \ (\gamma-1)\, n^{-\frac{1}{\gamma-1}} \int_{n^{\frac{\gamma}{\gamma-1}}\lambda}^\infty d\mu\, e^{-\gamma \mu^{-\frac{\gamma-1}{\gamma}}}\,.$$

Thus, by Lemma 3.22, for all $\lambda \in (0,\infty)$ and for all sufficiently large n, $\mathcal{L}_\lambda(h_n^+)$ and $\mathcal{L}_\lambda(h_n^-)$ are bounded by $A\lambda^{\frac{1}{\gamma}} \exp(-\gamma n \lambda^{\frac{\gamma-1}{\gamma}})$, where A is a positive constant. Thus,

$$\forall \lambda \in [\lambda_1, \infty), \quad \sum_{n\geq 0} |\beta_n| \big(\mathcal{L}_\lambda(h_n^+) + \mathcal{L}_\lambda(h_n^-)\big) \leq 2A \sum_{n\geq 1} |\beta_n| \lambda^{\frac{1}{\gamma}} e^{-\gamma n \lambda^{\frac{\gamma-1}{\gamma}}} < \infty, \tag{3.161}$$

the last inequality being a consequence of (3.159).

Next, deduce from (3.141) in Lemma 3.21 that for all fixed $x \in (0,\infty)$ and for all sufficiently large n,

$$\sup_{y\in[0,x]} h_n^+ \quad \text{and} \quad \sup_{y\in[0,x]} h_n^- \quad \leq \quad B n^q x^{-\frac{\gamma+3}{2}} \exp\big(-(\gamma-1)^{\gamma-1} n^\gamma x^{-(\gamma-1)}\big),$$

where $q = \frac{\gamma(\gamma+3)}{2(\gamma-1)} - \frac{\gamma+1}{\gamma-1}$ and where B is a positive constant only depending on γ. Since $\gamma > 1$, $n^\gamma \geq n$; this combined with (3.159) entails that for all $x \in \mathbb{R}_+$,

$$\sum_{n\geq 1} |\beta_n| \big(\sup_{y\in[0,x]} h_n^+ + \sup_{y\in[0,x]} h_n^- \big) < \infty \, . \tag{3.162}$$

By (3.160), (3.161) and (3.162), Lemma 3.18 applies and we get

$$\forall x \in \mathbb{R}_+, \quad f_\Gamma(x) = c_\gamma x^{-1-\frac{1}{\gamma}} \mathbf{N}_{\mathrm{nr}}\big(\Gamma > x^{-\frac{\gamma-1}{\gamma}}\big) = \sum_{n\geq 1} \beta_n \theta_n(x) \, .$$

This proves

$$\forall r \in (0,\infty), \quad c_\gamma \mathbf{N}_{\mathrm{nr}}\big(\Gamma > r\big) = \sum_{n\geq 1} \beta_n \, (nr)^{-\frac{\gamma+1}{\gamma-1}} \theta\big((nr)^{-\frac{\gamma}{\gamma-1}}\big), \tag{3.163}$$

which implies (3.60). Note that (3.162) and (3.150) with $x_1 = e^{-\gamma y_1}$ in Lemma 3.24 imply (3.59) in Theorem 3.5.

It remains to prove the asymptotic expansion (3.61). To that end, recall that $\xi(r) = r^{-\frac{\gamma+1}{\gamma-1}} \theta\big(r^{-\frac{\gamma}{\gamma-1}}\big)$, for all $r \in \mathbb{R}_+$. Then (3.57) in Proposition 3.4 easily entails that for any integer $N \geq 1$, as $r \to \infty$,

$$\frac{1}{C_1^*} r^{-1-\frac{\gamma}{2}} e^{r^\gamma} \xi\big(r(\gamma-1)^{-\frac{\gamma-1}{\gamma}}\big) \; = \; 1 + \sum_{1\leq n<N} V_n \, r^{-n\gamma} \; + \mathcal{O}_{N,\gamma}\big(r^{-N\gamma}\big) \; , \tag{3.164}$$

where $C_1^* := (2\pi)^{-\frac{1}{2}} (\gamma-1)^{\frac{1}{2}+\frac{1}{\gamma}} \gamma^{\frac{1}{2}}$ and where the sequence $(V_n)_{n\geq 1}$ is recursively defined by (3.56) in Proposition 3.4. This first implies that there exists $A, r_1 \in (0,\infty)$ that only depend on γ such that

$$\forall r \in (r_1,\infty), \; \forall n \geq 2, \quad \big|\xi\big(nr(\gamma-1)^{-\frac{\gamma-1}{\gamma}}\big)\big| \leq A r^{1+\frac{\gamma}{2}} e^{-n2^{\gamma-1} r^\gamma} \, . \tag{3.165}$$

Recall from Proposition 3.4 that there exists $x_1 \in (0,1)$ such that $\sum_{n\geq 1} |\beta_n| x_1^n < \infty$. Without loss of generality, we can choose r_1 such that $\exp(-2^{\gamma-1} r_1^\gamma) \leq x_1$. Then (3.163) and (3.165) imply that

$$\mathbf{N}_{\mathrm{nr}}\Big(\Gamma > r(\gamma-1)^{-\frac{\gamma-1}{\gamma}}\Big) = c_\gamma^{-1} \beta_1 \, \xi\big(r(\gamma-1)^{-\frac{\gamma-1}{\gamma}}\big) + \mathcal{O}_\gamma\big(r^{1+\frac{\gamma}{2}} e^{-2^\gamma r^\gamma}\big), \quad \text{as } r \to \infty,$$

and (3.164) implies (3.61) since $C_1 = c_\gamma^{-1} \beta_1 C_1^*$, where we recall from (3.41) that $c_\gamma^{-1} = \gamma \Gamma_e\big(\frac{\gamma-1}{\gamma}\big)$ and where we recall from Lemma 3.24 that $\beta_1 = \exp(C_0)$. This completes the proof of Theorem 3.5.

3.4.3 Proof of Theorem 3.7.

We first set

$$\forall x \in (0,\infty), \quad f_D(x) := c_\gamma x^{-1-\frac{1}{\gamma}} \mathbf{N}_{\mathrm{nr}}\big(D > 2x^{-\frac{\gamma-1}{\gamma}}\big) \, . \tag{3.166}$$

Then, Proposition 3.3, (3.46) and (3.47) imply for all $\lambda \in (0,\infty)$,

$$\mathcal{L}_\lambda(f_D) = \int_0^\infty dx \, e^{-\lambda x} f_D(x) = \mathrm{L}_\lambda(1,0) = \lambda^{\frac{1}{\gamma}} \mathrm{L}_1\big(\lambda^{\frac{\gamma-1}{\gamma}}, 0\big) \, . \tag{3.167}$$

We next use Lemma 3.25: let λ_2 be such that $\lambda_2^{\frac{\gamma-1}{\gamma}} = y_2$; then the sequences $(\gamma_n)_{n\geq 0}$ and $(\delta_n)_{n\geq 0}$ satisfy

$$\forall \lambda \in [\lambda_2, \infty), \quad \sum_{n\geq 0} \big(n|\gamma_n|\lambda^{\frac{\gamma-1}{\gamma}} + |\delta_n|\big)\lambda^{\frac{1}{\gamma}} e^{-\gamma n \lambda^{\frac{\gamma-1}{\gamma}}} < \infty$$

and

$$\mathcal{L}_\lambda(f_D) = \sum_{n\geq 0} n\gamma_n \lambda e^{-\gamma n\lambda^{\frac{\gamma-1}{\gamma}}} + \sum_{n\geq 0} \delta_n \lambda^{\frac{1}{\gamma}} e^{-\gamma n \lambda^{\frac{\gamma-1}{\gamma}}} \tag{3.168}$$

Recall from (3.54) in Proposition 3.4 the definition of θ and recall Proposition 3.6 that provides properties of the derivative s'_γ of the density s_γ given by (3.52). For all $n\geq 2$, and all $x\in(0,\infty)$, we set

$$\overline{\theta}_n(x) = n^{-\frac{2\gamma}{\gamma-1}} s'_\gamma\big(n^{-\frac{\gamma}{\gamma-1}}x\big) \quad \text{and} \quad \theta_n(x) = n^{-\frac{\gamma+1}{\gamma-1}}\theta\big(n^{-\frac{\gamma}{\gamma-1}}x\big).$$

Then, Proposition 3.6 and Proposition 3.4 imply that $\overline{\theta}_n$ and θ_n are continuous and Lebesgue integrable, and that

$$\forall \lambda \in \mathbb{R}_+, \quad \mathcal{L}_\lambda(\overline{\theta}_n) = \lambda e^{-\gamma n \lambda^{\frac{\gamma-1}{\gamma}}} \quad \text{and} \quad \mathcal{L}_\lambda(\theta_n) = \lambda^{\frac{1}{\gamma}} e^{-\gamma n\lambda^{\frac{\gamma-1}{\gamma}}} .$$

Thus,

$$\forall \lambda \in \mathbb{R}_+, \quad \mathcal{L}_\lambda(f_D) = \sum_{n\geq 2} \mathcal{L}_\lambda\big(n\gamma_n\overline{\theta}_n + \delta_n\theta_n\big).$$

We argue as in the proof of Theorem 3.5 using Lemma 3.18 to deduce that

$$\forall x \in \mathbb{R}_+, \quad f_D(x) = c_\gamma x^{-1-\frac{1}{\gamma}} \mathbf{N}_{\mathrm{nr}}\big(D > 2x^{-\frac{\gamma-1}{\gamma}}\big) = \sum_{n\geq 1} n\gamma_n \overline{\theta}_n(x) + \delta_n\theta_n(x) ,$$

the sum of functions being normally convergent on every compact subset of $\mathbb{R}_+$. This easily entails that

$$\forall r \in (0,\infty), \quad c_\gamma \mathbf{N}_{\mathrm{nr}}\big(D > 2r\big) = \sum_{n\geq 1} \gamma_n\, (nr)^{-\frac{\gamma+1}{\gamma-1}} s'_\gamma\big((nr)^{-\frac{\gamma}{\gamma-1}}\big) + \delta_n\, (nr)^{-\frac{\gamma+1}{\gamma-1}}\theta\big((nr)^{-\frac{\gamma}{\gamma-1}}\big), \tag{3.169}$$

which is (3.68). Note that (3.67) is an easy consequence of the estimate (3.65) in Proposition 3.6, of (3.57) in Proposition 3.4 and of Lemma 3.25 with $x_2 = e^{-\gamma y_2}$. Recall from (3.66) and (3.58) the following notation,

$$\forall r \in \mathbb{R}_+, \qquad \overline{\xi}(r) = r^{-\frac{\gamma+1}{\gamma-1}} s'_\gamma\big(r^{-\frac{\gamma}{\gamma-1}}\big) \quad \text{and} \quad \xi(r) = r^{-\frac{\gamma+1}{\gamma-1}}\theta\big(r^{-\frac{\gamma}{\gamma-1}}\big) .$$

Note that (3.68) implies

$$c_\gamma \mathbf{N}_{\mathrm{nr}}\big(D > r\big) = \gamma_2\overline{\xi}(r) + \delta_2\xi(r) + \sum_{n\geq 3} \gamma_n\overline{\xi}(nr/2) + \delta_n\xi(nr/2) . \tag{3.170}$$

Then, recall from (3.164) the asymptotic expansion of ξ and deduce from (3.65) in Proposition 3.6 that

$$\frac{1}{C_1^*} r^{-1-\frac{3\gamma}{2}} e^{r^\gamma}\, \overline{\xi}\big(r(\gamma-1)^{-\frac{\gamma-1}{\gamma}}\big) \; = \; 1 + \sum_{1\leq n<N} T_n\, r^{-n\gamma} \; + \mathcal{O}_{N,\gamma}\big(r^{-N\gamma}\big) , \tag{3.171}$$

where $C_1^* := (2\pi)^{-\frac{1}{2}}(\gamma-1)^{\frac{1}{2}+\frac{1}{\gamma}}\gamma^{\frac{1}{2}}$ and where the sequence $(T_n)_{n\geq 1}$ is recursively defined by (3.64) in Proposition 3.6. We easily deduce from the asymptotic expansions (3.164) and (3.171) that there exists $B, r_2 \in (0,\infty)$ such that for all $r\in(r_2,\infty)$ and for all $n\geq 3$,

$$\big|\overline{\xi}\,\big(\tfrac{1}{2}nr(\gamma-1)^{-\frac{\gamma-1}{\gamma}}\big)\big| \text{ and } \big|\xi\,\big(\tfrac{1}{2}nr(\gamma-1)^{-\frac{\gamma-1}{\gamma}}\big)\big| \;\leq B r^{1+\frac{3\gamma}{2}} e^{-n3^{\gamma-1}2^{-\gamma}r^\gamma} . \tag{3.172}$$

This combined with (3.170) implies that

$$\mathbf{N}_{\mathrm{nr}}\big(D>r(\gamma-1)^{-\frac{\gamma-1}{\gamma}}\big) = c_\gamma^{-1}\gamma_2\overline{\xi}\big(r(\gamma-1)^{-\frac{\gamma-1}{\gamma}}\big) + c_\gamma^{-1}\delta_2\xi\big(r(\gamma-1)^{-\frac{\gamma-1}{\gamma}}\big) + \mathcal{O}_\gamma\big(r^{1+\frac{3\gamma}{2}}e^{-n(3/2)^\gamma r^\gamma}\big),$$

as $r\to\infty$. Then (3.164), by (3.171) imply

$$\begin{aligned}\mathbf{N}_{\mathrm{nr}}\big(D>r(\gamma-1)^{-\frac{\gamma-1}{\gamma}}\big) =& c_\gamma^{-1}\gamma_2 C_1^* r^{1+\frac{3\gamma}{2}}e^{-r^\gamma}\\ &+\sum_{1\le n<N} c_\gamma^{-1}C_1^*(\gamma_2 T_n+\delta_2 V_{n-1})r^{-n\gamma+1+\frac{3\gamma}{2}}e^{-r^\gamma} + \mathcal{O}_{N,\gamma}\big(r^{-N\gamma+1+\frac{3\gamma}{2}}e^{-r^\gamma}\big)\end{aligned} \tag{3.173}$$

Recall from (3.151) in Lemma 3.25 that $\gamma_2=\frac{1}{2}\gamma(\gamma-1)e^{2C_0}$ and $\delta_2=-\frac{1}{2}(\gamma+1)e^{2C_0}$. This implies (3.69) with

$$C_2=c_\gamma^{-1}C_1^*\gamma_2 \quad\text{and}\quad \forall n\ge 1,\quad U_n=T_n+\tfrac{\delta_2}{\gamma_2}V_{n-1}=T_n-\tfrac{\gamma+1}{\gamma(\gamma-1)}V_{n-1}\ .$$

This completes the proof of Theorem 3.7.

3.5 Appendix: proof of Lemma 3.9.

We first recall the following notation from Introduction: let $h\in\mathbb{C}(\mathbb{R}_+,\mathbb{R}_+)$. For any $a\in[0,h(0)]$, set

$$\ell_a(h)=\inf\big\{t\in\mathbb{R}_+ : h(t)=h(0)-a\big\}\quad\text{and}\quad r_a(h)=\inf\big\{t\in(0,\infty) : h(0)-a>h(t)\big\}\ , \tag{3.174}$$

with the convention that $\inf\emptyset=\infty$. Standard results on stopping times assert that $\ell_a(h)$ and $r_a(h)$ are $[0,\infty]$-valued Borel measurable functions of h: see for instance Revuz & Yor [97], Chapter I, Proposition 4.5 and Proposition 4.6, p. 43. Moreover, it is easy to check that for a fixed h, $a\mapsto\ell_a(h)$ is left continuous and that $a\mapsto r_a(h)$ is right continuous. By standard arguments, $(a,h)\mapsto(\ell_a(h),r_a(h))$ is Borel measurable on the set $A:=\{(a,h)\in\mathbb{R}_+\times\mathbb{C}(\mathbb{R}_+,\mathbb{R}_+) : a\le h(0)\}$. We next recall the following notation: for all $(a,h)\in A$, we set

$$\forall s\in\mathbb{R}_+,\quad \mathcal{E}_s(h,a):=h\big((\ell_a(h)+s)\wedge r_a(h)\big)-h(0)+a\ ,$$

with the convention that $\mathcal{E}(h,a)$ is the null function $\mathbf{0}$ is $\ell_a(h)=\infty$. The previous arguments entail that

$$(a,h)\in A\mapsto\mathcal{E}(h,a)\in\mathbb{C}(\mathbb{R}_+,\mathbb{R}_+)\text{ is Borel measurable.} \tag{3.175}$$

Recall from (3.73) the definition of $\mathtt{Exc}$. Recall that $p_H:[0,\zeta_H]\to\mathcal{T}_H$ stands for the canonical projection and recall from (3.6) that the mass measure $\mathbf{m}_H$ is the pushforward measure of the Lebesgue measure on $[0,\zeta_H]$ by p_H. Suppose that there exist $r,s\in(0,\zeta_H)$ such that $r<s$ and such that H is constant on (r,s). Thus $p_H((r,s))=\{p_H(r)\}$ and $m_H(\{p_H(r)\})\ge s-r>0$, which contradicts the fact that $\mathbf{m}_H$ is diffuse. Recall from (3.5) the definition of the set of leaves $\mathtt{Lf}(\mathcal{T}_H)$ of $\mathcal{T}_H$. Suppose there exist $r,s\in(0,\zeta_H)$ such that $r<s$ and such that H is strictly monotone on (r,s). It easily implies that $p_H((r,s))\subset\mathcal{T}_H\backslash\mathtt{Lf}(\mathcal{T}_H)$, but $\mathbf{m}_H(p_H((r,s)))\ge s-r>0$, which contradicts the fact that $\mathbf{m}_H\big(\mathcal{T}_H\backslash\mathtt{Lf}(\mathcal{T}_H)\big)=0$. Thus, we have proved the following.

($*$) *Let $H\in\mathtt{Exc}$. Let $r,s\in(0,\zeta_H)$ be such that $r<s$. Then on (r,s), H is neither non-increasing nor non-decreasing.*

Let $t\in(0,\infty)$ and $H\in\mathtt{Exc}$ be such that $\zeta_H>t$. Recall the following notation

$$\forall s\in\mathbb{R}_+,\quad H_s^-=H_{(t-s)_+},\quad H_s^+=H_{t+s},\quad \overleftarrow{H}^a:=\mathcal{E}(H^-,a)\quad\text{and}\quad \overrightarrow{H}^a:=\mathcal{E}(H^+,a),$$

for all $a \in [0, H_t]$. Note that $H_0^- = H_0^+ = H_t$. We also recall the following notation

$$\mathcal{M}_{0,t}(H) = \sum_{a \in \mathcal{J}_{0,t}} \delta_{(a, \overleftarrow{H}^a, \overrightarrow{H}^a)} , \tag{3.176}$$

where $\mathcal{J}_{0,t} := \big\{ a \in [0, H_t] : \text{either } \ell_a(H^-) < r_a(H^-) \text{ or } \ell_a(H^+) < r_a(H^+) \big\}$, which is countable. Then, the definitions (3.174) and ($*$) entail that

$$\forall t \in (0, \infty),\ \forall H \in \mathtt{Exc} \text{ such that } \zeta_H > t, \quad \text{the closure of } \mathcal{J}_{0,t} \text{ is } [0, H_t]. \tag{3.177}$$

We next introduce the compact set $C_t := \{ s \in [0, \zeta_H - t] : H_{t+s} = \inf_{r \in [t,t+s]} H_r \}$, whose Lebesgue measure is denoted by $|C_t|$. We easily check that $p_H(C_t) \subset \{\rho, p_H(t)\} \cup \big(\mathcal{T}_H \backslash \mathtt{Lf}(\mathcal{T}_H)\big)$. Since $\mathbf{m}_H$ is diffuse and supported by the set of leaves of $\mathcal{T}_H$, we get $0 = \mathbf{m}_H(p_H(C_t)) \geq |C_t|$, which implies that $|C_t| = 0$. Then note that for all $a \in [0, H_t]$,

$$[0, \ell_a(H^+)] \backslash C_t \subset \big\{ s \in [0, \ell_a(H^+)] : H_{t+s} > \inf_{r \in [t,t+s]} H_r \big\} \subset \bigcup_{b \in \mathcal{J}_{0,t} \cap [0,a)} \big(\ell_b(H^+), r_b(H^+) \big) \subset [0, \ell_a(H^+)].$$

Since $|C_t| = 0$, this entails,

$$\forall a \in [0, H_t], \quad \ell_a(H^+) = \sum_{b \in \mathcal{J}_{0,t}} \mathbf{1}_{[0,a)}(b) \big(r_b(H^+) - \ell_b(H^+) \big) = \sum_{b \in \mathcal{J}_{0,t}} \mathbf{1}_{[0,a)}(b) \zeta_{\overrightarrow{H}^b} .$$

Similar arguments imply that

$$\begin{aligned} \forall a \in [0, H_t], \quad \ell_a(H^+) &= \sum_{b \in \mathcal{J}_{0,t}} \mathbf{1}_{[0,a)}(b) \zeta_{\overrightarrow{H}^b}, \quad \ell_a(H^-) = \sum_{b \in \mathcal{J}_{0,t}} \mathbf{1}_{[0,a)}(b) \zeta_{\overleftarrow{H}^b}, \\ r_a(H^+) &= \sum_{b \in \mathcal{J}_{0,t}} \mathbf{1}_{[0,a]}(b) \zeta_{\overrightarrow{H}^b}, \quad r_a(H^-) = \sum_{b \in \mathcal{J}_{0,t}} \mathbf{1}_{[0,a]}(b) \zeta_{\overleftarrow{H}^b}. \end{aligned} \tag{3.178}$$

Moreover, since H is continuous with compact support, we immediately get

$$\forall \varepsilon, \eta \in (0, \infty), \quad \# \big\{ a \in \mathcal{J} : \Gamma(\overleftarrow{H}^a) \vee \Gamma(\overrightarrow{H}^a) > \eta \quad \text{or} \quad \zeta_{\overleftarrow{H}^a} \vee \zeta_{\overrightarrow{H}^a} > \varepsilon \big\} < \infty . \tag{3.179}$$

We next easily see that $\mathcal{T}_{\overrightarrow{H}^a}$ can be identified with a subtree of $\mathcal{T}_H$; therefore, up to this identification, the set of leaves of $\mathcal{T}_{\overrightarrow{H}^a}$ is contained in the set of leaves of $\mathcal{T}_H$ and $\mathbf{m}_{\overrightarrow{H}^a}$ is the restriction of $\mathbf{m}_H$ to $\mathcal{T}_{\overrightarrow{H}^a}$. This implies that $\mathbf{m}_{\overrightarrow{H}^a}$ is diffuse and supported by the set of leaves of $\mathcal{T}_{\overrightarrow{H}^a}$. Namely, $\overrightarrow{H}^a \in \mathtt{Exc}$. A similar argument show that $\overleftarrow{H}^a \in \mathtt{Exc}$. This fact combined with (3.177) and (3.179) imply the following:

$$\forall t \in (0, \infty),\ \forall H \in \mathtt{Exc} \text{ such that } \zeta_H > t, \quad \mathcal{M}_{0,t}(H) \in \mathscr{M}_{\mathrm{pt}}(E) , \tag{3.180}$$

where $\mathscr{M}_{\mathrm{pt}}(E)$ is as in Definition 3.2.1. Moreover (3.175) easily implies that $(a, t, H) \mapsto (\overleftarrow{H}^a, \overrightarrow{H}^a)$ is Borel-measurable, which immediately implies Lemma 3.9 (i).

Let us prove Lemma 3.9 (ii). Recall from Definition 3.2.1 the definition of the sigma field $\mathcal{G}$ on $\mathscr{M}_{\mathrm{pt}}(E)$. We next fix $t \in (0, \infty)$ and $H \in \mathtt{Exc}$ such that $\zeta_H > t$. First note that (3.178) imply that $\ell_a(H^+)$ and $r_a(H^+)$ are $\mathscr{B}(\mathbb{R}_+) \otimes \mathcal{G}$-measurable functions of $(a, \mathcal{M}_{0,t}(H))$, where $\mathscr{B}(\mathbb{R}_+)$ stands for the Borel sigma field on $\mathbb{R}_+$. We then fix $s \in \mathbb{R}_+$ and we set $a(s) = \inf\{ a \in \mathbb{R}_+ : r_a(H^+) > s \}$, with the convention that $\inf \emptyset = \infty$. The previous argument and the fact that $a \mapsto r_a(H^+)$ is right continuous entail that $a(s)$ can be viewed as a $\mathcal{G}$-measurable function of $\mathcal{M}_{0,t}(H)$. Then note that if $a(s) < \infty$, then

$$H_{t+s} = H_s^+ = H_t - a(s) + \overrightarrow{H}^{a(s)} \big(s - \ell_{a(s)}(H^+) \big) . \tag{3.181}$$

Then for all $a \in \mathbb{R}_+$, we set $N_a = \sum_{b \in \mathcal{J}_{0,t}} \mathbf{1}_{(a,\infty)}(b) \mathbf{1}_{\{\zeta_{\overrightarrow{H}^b} > 0\}}$. We have actually have proved previously that the closure of the set $\{b \in \mathcal{J}_{0,t} : \ell_b(H^+) < r_b(H^+)\}$ is $[0, H_t]$. Thus $H_t = \inf\{a \in \mathbb{R}_+ : N_a > 0\}$, which proves that H_t is a $\mathcal{G}$-measurable function of $\mathcal{M}_{0,t}(H)$. Moreover $(a, \mathcal{M}_{0,t}(H)) \mapsto \overrightarrow{H}^a$ is $\mathscr{B}(\mathbb{R}_+) \otimes \mathcal{G}$-measurable. Thus, (3.181) implies that H_s^+ is a $\mathcal{G}$-measurable function of $\mathcal{M}_{0,t}(H)$. Since the Borel sigma field on $\mathbb{C}(\mathbb{R}_+, \mathbb{R}_+)$ is generated by coordinate applications, this implies that H^+ is a $\mathcal{G}$-measurable function of $\mathcal{M}_{0,t}(H)$. A similar argument shows that H^- is also a $\mathcal{G}$-measurable function of $\mathcal{M}_{0,t}(H)$, which easily completes the proof of Lemma 3.9 (ii). ■

Chapter 4

Cutting down p-trees and inhomogeneous continuum random trees

The results of this chapter are from the joint work [39] with Nicolas Broutin, submitted for publication.

Contents

We study a fragmentation of birthday trees, and give exact correspondences between the birthday trees and the trees which encode the fragmentation. We then use these results to study the fragmentation of the ICRTs (scaling limits of $\boldsymbol{p}$-trees) and give distributional correspondences between the ICRT and the tree encoding the fragmentation. The results for the ICRT extend the results of Bertoin and Miermont [30] about the cut tree of the Brownian continuum random tree.

4.1 Introduction

The study of random cutting of trees has been initiated by Meir and Moon [88] in the following form: Given a (graph theoretic) tree, one can proceed to chop the tree into pieces by iterating the following process: choose a uniformly random edge; removing it disconnects the tree into two pieces; discard the part which does not contain the root and keep chopping the portion containing the root until it is reduced to a single node. In the present document, we consider the related version where the vertices are chosen at random and removed (until one is left with an empty tree); each such pick is referred to as a *cut*. We will see that this version is actually much more adapted than the edge cutting procedure to the problems we consider here.

The main focus in [88] and in most of the subsequential papers has been put on the study of some parameters of this cutting down process, and in particular on how many cuts are necessary for the process to finish. This has been studied for a number of different models of deterministic and random trees such as complete binary trees of a given height, random trees arising from the divide-and-conquer paradigm [48, 67, 68, 70] and the family trees of finite-variance critical Galton–Watson processes conditioned on the total progeny [60, 72, 93]. The latter model of random trees turns out to be far more interesting, and it provides an *a posteriori* motivation for the cutting down process. As we will see shortly, the cutting down process provides an interesting way to investigate some of the structural properties of random trees by partial destruction and re-combination, or equivalently as partially resampling the tree.

Let us now be more specific: if L_n denotes the number of cuts required to completely cut down a uniformly labelled rooted tree (random Cayley tree, or equivalently condition Galton–Watson tree with Poisson offspring distribution) on n nodes, then $n^{-1/2}L_n$ converges in distribution to a Rayleigh distribution which has density $xe^{-x^2/2}$ on $\mathbb{R}_+$. Janson [72] proved that a similar result holds for any Galton–Watson tree with a finite-variance offspring distribution conditioned on the total progeny to be n. This is the parameter point of view. Addario-Berry, Broutin, and Holmgren [7] have shown that for the random Cayley trees, L_n actually has the same distribution as the number of nodes on the path between two uniformly random nodes. Their method relies on an "objective" argument based on a coupling that associates with the cutting procedure a partial resampling of the Cayley tree of the kind mentioned earlier: if one considers the (ordered) sequence of subtrees which are discarded as the cutting process goes on, and adds a path linking their roots, then the resulting tree is a uniformly random Cayley tree, and the two extremities of the path are independent uniform random nodes. So the properties of the parameter L_n follow from a stronger correspondence between the combinatorial objects themselves.

This strong connection between the discrete objects can be carried to the level of their scaling limit, namely Aldous' Brownian continuum random tree (CRT) [10]. Without being too precise for now, the natural cutting procedure on the Brownian CRT involves a Poisson rain of cuts sampled according to the length measure. However, not all the cuts contribute to the isolation of the root. As in the partial resampling of the discrete setting, we glue the sequence of discarded subtrees along an interval, thereby obtaining a new CRT. If the length of the interval is well-chosen (as a function of the cutting process), the tree obtained is distributed like the Brownian CRT and the two ends of the interval are independently random leaves. This identifies the distribution of the discarded subtrees from the cutting procedure as the distribution of the forest one obtains from a spinal decomposition of the Brownian CRT. The distribution of the latter is intimately related to Bismut's [33] decomposition of a Brownian excursion. See also [52] for the generalization to the Lévy case. Note that a similar identity has been proved by Abraham and Delmas [5] for general Lévy trees without using a discrete approximation. A related example is that of the subtree prune and re-graft dynamics of Evans et al. [59] [See also 57], which is even closer to the cutting procedure and truly resamples the object rather than giving a "recursive" decomposition.

The aim of this work is two-fold. First we prove exact identities and give reversible transformations of $\boldsymbol{p}$-trees similar to the ones for Cayley trees in [7]. The model of $\boldsymbol{p}$-trees introduced by Camarri and

Pitman [41] generalizes Cayley trees in allowing "weights" on the vertices. In particular, this additional structure of weights introduces some inhomogeneity. We then lift these results to the scaling limits, the inhomogeneous continuum random trees (ICRT) of Aldous and Pitman [13], which are closely related to the general additive coalescent [13, 23, 24]. Unlike the Brownian CRT or the stable trees (special cases of Lévy trees), a general ICRT is not self-similar. Nor does it enjoy a "branching property" as the Lévy trees do [83]. This lack of "recursivity" ruins the natural approaches such as the one used in [4, 5] or the ones which would argue by comparing two fragmentations with the same dislocation measure but different indices of self-similarity [25]. This is one of the reasons why we believe these path transformations at the level of the ICRT are interesting. Furthermore, a conjecture of Aldous, Miermont, and Pitman [15, p. 185] suggests that the path transformations for ICRTs actually explain the result of Abraham and Delmas [5] for Lévy trees by providing a result "conditional on the degree distribution".

Second, rather than only focusing on the isolation of the root we also consider the genealogy of the entire fragmentation as in the recent work of Bertoin and Miermont [30] and Dieuleveut [46] (who examine the case of Galton–Watson trees). In some sense, this consists in obtaining transformations corresponding to tracking the effect of the cutting down procedure on the isolation of all the points simultaneously. Tracking finitely many points is a simple generalization of the one-point results, but the "complete" result requires additional insight. The results of the present document are used in Chapter 5 to prove that the "complete" cutting procedure in which one tries to isolate every point yields a construction of the genealogy of the fragmentation on ICRTs which is reversible in the case of the Brownian CRT. More precisely, the genealogy of Aldous–Pitman's fragmentation of a Brownian CRT is another Brownian CRT, say $\mathcal{G}$, and there exists a random transformation of $\mathcal{G}$ into a real tree $\mathcal{T}$ such that in the pair $(\mathcal{T}, \mathcal{G})$ the tree $\mathcal{G}$ is indeed distributed as the genealogy of the fragmentation on $\mathcal{T}$, conditional on $\mathcal{T}$. The proof there relies crucially on the "bijective" approach that we develop here.

Plan of the chapter. In the next section, we introduce the necessary notation and relevant background. We then present more formally the discrete and continuous models we are considering, and in which sense the inhomogeneous continuum random trees are the scaling limit of $\boldsymbol{p}$-trees. In Section 4.3 we introduce the cutting down procedures and state our main results. The study of cutting down procedure for $\boldsymbol{p}$-trees is the topic of Section 4.4. The results are lifted to the level of the scaling limits in Section 4.5.

4.2 Notation, models and preliminaries

Although we would like to introduce our results earlier, a fair bit of notation and background is in order before we can do so properly. This section may safely be skipped by the impatient reader and referred to later on.

4.2.1 Aldous–Broder Algorithm and p-trees

Let A be a finite set and $\boldsymbol{p} = (p_u, u \in A)$ be a probability measure on A such that $\min_{u\in A} p_u > 0$; this ensures that A is indeed the support of $\boldsymbol{p}$. Let $\mathbb{T}_A$ denote the set of rooted trees labelled with (all the) elements of A (connected acyclic graphs on A, with a distinguished vertex). For $t \in \mathbb{T}_A$, we let $r = r(t)$ denote its root vertex. For $u, v \in A$, we write $\{u, v\}$ to mean that u and v are adjacent in t. We sometimes write $\langle u, v\rangle$ to mean that $\{u, v\}$ is an edge of t, and that u is on the path between r and v (we think of the edges as pointing towards the root). For a tree $t \in \mathbb{T}_A$ (rooted at r, say) and a node $v \in A$, we let t^v denote the tree re-rooted at v.

We usually abuse notation, but we believe it does not affect the clarity or precision of our statements. For instance, we refer to a node u in the vertex set $\mathfrak{v}(t)$ of a tree t using $u \in t$. Depending on the context, we sometimes write $t \setminus \{u\}$ to denote the forest induced by t on the vertex set $\mathfrak{v}(t) \setminus \{u\}$. The (in-)degree $C_u(t)$ of a vertex $u \in A$ is the number of edges of the form $\langle u, v\rangle$ with $v \in A$. For a rooted tree t, and a node u of t, we write $\mathrm{Sub}(t, u)$ for the subtree of t rooted at u (above u). For $t \in \mathbb{T}_A$ and $\mathbf{V} \subseteq A$, we write $\mathrm{Span}(t; \mathbf{V})$ for the subtree of t spanning $\mathbf{V}$ and the root of $r(t)$. So $\mathrm{Span}(t; V)$ is the subtree induced by t on the set

$$\bigcup_{u \in \mathbf{V}} [\![r(t), u]\!],$$

where $[\![u, v]\!]$ denotes collection of nodes on the (unique) path between u and v in t. When $\mathbf{V} = \{v_1, v_2, \dots, v_k\}$, we usually write $\mathrm{Span}(t; v_1, \dots, v_k)$ instead of $\mathrm{Span}(t; \{v_1, \dots, v_k\})$. We also write

$$\mathrm{Span}^*(t; \mathbf{V}) := \mathrm{Span}(t; \mathbf{V}) \setminus \{r(t)\}.$$

As noticed by Aldous [18] and Broder [37], one can generate random trees on A by extracting a tree from the trace of a random walk on A, where the sequence of steps is given by a sequence of i.i.d. vertices distributed according to $\boldsymbol{p}$.

Algorithm (Weighted version of Aldous–Broder Algorithm)**.** Let $\mathbf{Y} = (Y_j, j \geq 0)$ be a sequence of independent variables with common distribution $\boldsymbol{p}$; further on, we say that Y_j are i.i.d. $\boldsymbol{p}$-nodes. Let $\mathcal{T}(\mathbf{Y})$ be the graph rooted at Y_0 with the set of edges

$$\{\langle Y_{j-1}, Y_j \rangle : Y_j \notin \{Y_0, \cdots, Y_{j-1}\}, j \geq 1\}. \tag{4.1}$$

The sequence $\mathbf{Y}$ defines a random walk on A, which eventually visits every element of A with probability one, since A is the support of $\boldsymbol{p}$. So the trace $\{\langle Y_{j-1}, Y_j \rangle : j \geq 1\}$ of the random walk on A is a connected graph on A, rooted at Y_0. Algorithm 4.2.1 extracts the tree $\mathcal{T}(\mathbf{Y})$ from the trace of the random walk. To see that $\mathcal{T}(\mathbf{Y})$ is a tree, observe that the edge $\langle Y_{j-1}, Y_j \rangle$ is added only if Y_j has never appeared before in the sequence. It follows easily that $\mathcal{T}(\mathbf{Y})$ is a connected graph without cycles, hence a tree on A. Let π denote the distribution of $\mathcal{T}(\mathbf{Y})$.

Lemma 4.1 ([18, 37, 58])**.** *For $t \in \mathbb{T}_A$, we have*

$$\pi(t) := \pi^{(\boldsymbol{p})}(t) = \prod_{u \in A} p_u^{C_u(t)}. \tag{4.2}$$

Note that π is indeed a probability distribution on $\mathbb{T}_A$, since by Cayley's multinomial formula ([42, 96]), we have

$$\sum_{t \in \mathbb{T}_A} \pi(t) = \sum_{t \in \mathbb{T}_A} \prod_{u \in A} p_u^{C_u(t)} = \left(\sum_{u \in A} p_u \right)^{|A|-1} = 1. \tag{4.3}$$

A random tree on A distributed according to π as specified by (4.2) is called a $\boldsymbol{p}$*-tree*. It is also called the birthday tree in the literature, for its connection with the general birthday problem (see [41]). Observe that when $\boldsymbol{p}$ is the uniform distribution on $[n] := \{1, 2, \dots, n\}$, a $\boldsymbol{p}$-tree is a uniformly random rooted tree on $[n]$ (a Cayley tree). So the results we are about to present generalize the exact distributional results in [7]. However, we believe that the point of view we adopt here is a little cleaner, since it permits to make the transformation *exactly* reversible without any extra anchoring nodes (which prevent any kind duality at the discrete level).

From now on, we consider $n \geq 1$ and let $[n]$ denote the set $\{1, 2, \cdots, n\}$. We write $\mathbb{T}_n$ as a shorthand for $\mathbb{T}_{[n]}$, the set of the rooted trees on $[n]$. Let also $\boldsymbol{p} = (p_i, 1 \leq i \leq n)$ be a probability measure on $[n]$ satisfying $\min_{i \in [n]} p_i > 0$. For a subset $A \subseteq [n]$ such that $\boldsymbol{p}(A) > 0$, we let $\boldsymbol{p}|_A(\,\cdot\,) = \boldsymbol{p}(\,\cdot \cap A)/\boldsymbol{p}(A)$ denote the restriction of $\boldsymbol{p}$ on A, and write $\pi|_A := \pi^{(\boldsymbol{p}|_A)}$. The following lemma says that the distribution of $\boldsymbol{p}$-trees is invariant by re-rooting at an independent $\boldsymbol{p}$-node and "recursive" in a certain sense. These two properties are one of the keys to our results on the discrete objects. (For a probability distribution μ, we write $X \sim \mu$ to mean that μ is the distribution of the random variable X.)

Lemma 4.2. *Let T be a $\mathbf{p}$-tree on $[n]$.*

i) If V is an independent $\mathbf{p}$-node. Then, $T^V \sim \pi$.

ii) Let N be set of neighbors of the root in T. Then, for $u \in N$, conditional on $\mathfrak{v}(\mathrm{Sub}(T, u)) = \mathbf{V}$, $\mathrm{Sub}(T, u) \sim \pi|_{\mathbf{V}}$ independent of $\{\mathrm{Sub}(T, w) : w \in N, w \neq u\}$.

The first claim can be verified from (4.2), the second is clear from the product form of π.

4.2.2 Measured metric spaces and the Gromov–Prokhorov topology

If (X, d) is a metric space endowed with the Borel σ-algebra, we denote by $\mathcal{M}_f(X)$ the set of finite measures on X and by $\mathcal{M}_1(X)$ the subset of probability measures on X. If $m \in \mathcal{M}_f(X)$, we denote by $\mathrm{supp}(m)$ the support of m on X, that is the smallest closed set A such that $m(A^c) = 0$. If $f : X \to Y$ is a measurable map between two metric spaces, and if $m \in \mathcal{M}_f(X)$, then the push-forward of m is an element of $\mathcal{M}_f(Y)$, denoted by $f_* m \in \mathcal{M}_f(Y)$, and is defined by $(f_* m)(A) = m(f^{-1}(A))$ for each Borel set A of Y. If $m \in \mathcal{M}_f(X)$ and $A \subseteq X$, we denote by $m \restriction_A$ the restriction of m to A: $m \restriction_A (B) = m(A \cap B)$ for any Borel set B. This should not be confused with the restriction of a probability measure, which remains a probability measure and is denoted by $m|_A$.

We say a triple (X, d, μ) is a *measured metric space* (or sometimes a *metric measure space*) if (X, d) is a Polish space (separable and complete) and $\mu \in \mathcal{M}_1(X)$. Two measured metric spaces (X, d, μ) and (X', d', μ') are said to be *weakly isometric* if there exists an isometry ϕ between the supports of μ on X and of μ' on X' such that $(\phi)_* \mu = \mu'$. This defines an equivalence relation between the measured metric spaces, and we denote by $\mathbb{M}$ the set of equivalence classes. Note that if (X, d, μ) and (X', d', μ') are weakly isometric, the metric spaces (X, d) and (X', d') may not be isometric.

We can define a metric on $\mathbb{M}$ by adapting Prokhorov's distance. Consider a metric space (X, d) and for $\epsilon > 0$, let $A^\epsilon := \{x \in X : d(x, A) < \epsilon\}$. Then, given two (Borel) probability measures $\mu, \nu \in \mathcal{M}_1(X)$, the Prokhorov distance δ_{P} between μ and ν is defined by

$$\delta_{\mathrm{P}}(\mu, \nu) := \inf\{\epsilon > 0 : \mu(A) \leq \nu(A^\epsilon) + \epsilon \text{ and } \nu(A) \leq \mu(A^\epsilon) + \epsilon, \text{ for all Borel sets } A\}. \tag{4.4}$$

Note that the definition of the Prokhorov distance (4.4) can be easily extended to a pair of finite (Borel) measures on X. Then, for two measured metric spaces (X, d, μ) and (X', d', μ') the Gromov–Prokhorov (GP) distance between them is defined to be

$$\delta_{\mathrm{GP}}((X, d, \mu), (X', d', \mu')) = \inf_{Z, \phi, \psi} \delta_{\mathrm{P}}(\phi_* \mu, \psi_* \mu'),$$

where the infimum is taken over all metric spaces Z and isometric embeddings $\phi : \mathrm{supp}(\mu) \to Z$ and $\psi : \mathrm{supp}(\mu') \to Z$. It is clear that δ_{GP} depends only on the equivalence classes containing (X, d, μ) and (X', d', μ'). Moreover, the Gromov–Prokhorov distance turns $\mathbb{M}$ in a Polish space.

There is another more convenient characterization of the GP topology (the topology induced by δ_{GP}) that relies on convergence of distance matrices between random points. Let $\mathcal{X} = (X, d, \mu)$ be a measured metric space and let $(\xi_i, i \geq 1)$ be a sequence of i.i.d. points of common distribution μ. In the following, we will often refer to such a sequence as $(\xi_i, i \geq 1)$ as an i.i.d. μ-sequence. We write $\rho^{\mathcal{X}} = (d(\xi_i, \xi_j), 1 \leq i, j < \infty)$ for the distance matrix associated with this sequence. One easily verifies that the distribution of $\rho^{\mathcal{X}}$ does not depend on the particular element of an equivalent class of $\mathbb{M}$. Moreover, by Gromov's reconstruction theorem [64, $3\frac{1}{2}$], the distribution of $\rho^{\mathcal{X}}$ characterizes $\mathcal{X}$ as an element of $\mathbb{M}$.

Proposition 4.3 (Corollary 8 of [84])**.** *If $\mathcal{X}$ is some random element taking values in $\mathbb{M}$ and for each $n \geq 1$, $\mathcal{X}_n$ is a random element taking values in $\mathbb{M}$, then $\mathcal{X}_n$ converges to $\mathcal{X}$ in distribution as $n \to \infty$ if and only if $\rho^{\mathcal{X}_n}$ converges to $\rho^{\mathcal{X}}$ in the sense of finite-dimensional distributions.*

Pointed Gromov–Prokhorov topology. The above characterization by matrix of distances turns out to be quite handy when we want to keep track of marked points. Let $k \in \mathbb{N}$. If (X, d, μ) is a measured metric space and $\mathbf{x} = (x_1, x_2, \cdots, x_k) \in X^k$ is a k-tuple, then we say $(X, d, \mu, \mathbf{x})$ is a *k-pointed measured metric space*, or simply a pointed measured metric space. Two pointed metric measure spaces $(X, d, \mu, \mathbf{x})$ and $(X', d', \mu', \mathbf{x}')$ are said to be *weakly isometric* if there exists an isometric bijection

$$\phi : \mathrm{supp}(\mu) \cup \{x_1, x_2, \cdots, x_k\} \to \mathrm{supp}(\mu') \cup \{x'_1, x'_2, \cdots, x'_k\}$$

such that $(\phi)_*\mu = \mu'$ and $\phi(x_i) = x'_i$, $1 \le i \le k$, where $\mathbf{x} = (x_1, x_2, \cdots, x_k)$ and $\mathbf{x}' = (x'_1, x'_2, \cdots, x'_k)$. We denote by $\mathbb{M}^*_k$ the space of weak isometry-equivalence classes of k-pointed measured metric spaces. Again, we emphasize the fact that the underlying metric spaces (X, d) and (X', d') do not have to be isometric. The space $\mathbb{M}^*_k$ equipped with the following pointed Gromov–Prokhorov topology is a Polish space.

A sequence $(X_n, d_n, \mu_n, \mathbf{x}_n)_{n\ge1}$ of k-pointed measured metric spaces is said to converge to some pointed measured metric space $(X, d, \mu, \mathbf{x})$ in the k-pointed Gromov–Prokhorov topology if for any $m \ge 1$,

$$\big(d_n(\xi^*_{n,i}, \xi^*_{n,j}), 1 \le i, j \le m\big) \overset{n\to\infty}{\underset{d}{\longrightarrow}} \big(d(\xi^*_i, \xi^*_j), 1 \le i, j \le m\big),$$

where for each $n \ge 1$ and $1 \le i \le k$, $\xi^*_{n,i} = x_{n,i}$ if $\mathbf{x}_n = (x_{n,1}, x_{n,2}, \cdots, x_{n,k})$ and $(\xi^*_{n,i}, i \ge k+1)$ is a sequence of i.i.d. μ_n-points in X_n. Similarly, $\xi^*_i = x_i$ for $1 \le i \le k$ and $(\xi^*_i, i \ge k+1)$ is a sequence of i.i.d. μ-points in X. This induces the k-pointed Gromov–Prokhorov topology on $\mathbb{M}^*_k$.

4.2.3 Compact metric spaces and the Gromov–Hausdorff metric

Gromov–Hausdorff metric. Two compact subsets A and B of a given metric space (X, d) are compared using the Hausdorff distance δ_{H}.

$$\delta_{\mathrm{H}}(A, B) := \inf\{\epsilon > 0 : A \subseteq B^\epsilon \text{ and } B \subseteq A^\epsilon\}.$$

To compare two compact metric spaces (X, d) and (X', d'), we first embed them into a single metric space (Z, δ) via isometries $\phi : X \to Z$ and $\psi : X' \to Z$, and then compare the images $\phi(X)$ and $\psi(X')$ using the Hausdorff distance on Z. One then defines the Gromov–Hausdoff (GH) distance δ_{GH} by

$$\delta_{\mathrm{GH}}((X, d), (X', d')) := \inf_{Z,\phi,\psi} \delta_{\mathrm{H}}(\phi(X), \psi(X')),$$

where the infimum ranges over all choices of metric spaces Z and isometric embeddings $\phi : X \to Z$ and $\psi : X' \to Z$. Note that, as opposed to the case of the GP topology, two compact metric spaces that are at GH distance zero are isometric.

Gromov–Hausdorff–Prokhorov metric. Now if (X, d) and (X', d') are two compact metric spaces and if $\mu \in \mathcal{M}_f(X)$ and $\mu' \in \mathcal{M}_f(X')$, one way to compare simultaneously the metric spaces and the measures is to define

$$\delta_{\mathrm{GHP}}\left((X, d, \mu), (X', d', \mu')\right) := \inf_{Z,\phi,\psi} \Big\{ \delta_{\mathrm{H}}\left(\phi(X), \psi(X')\right) \vee \delta_{\mathrm{P}}(\phi_*\mu, \psi_*\mu') \Big\},$$

where the infimum ranges over all choices of metric spaces Z and isometric embeddings $\phi : X \to Z$ and $\psi : X' \to Z$. If we denote by $\mathbb{M}_c$ the set of equivalence classes of compact measured metric spaces under measure-preserving isometries, then $\mathbb{M}_c$ is Polish when endowed with δ_{GHP}.

Pointed Gromov–Hausdorff metric. We fix some $k \in \mathbb{N}$. Given two compact metric spaces (X, d_X) and (Y, d_Y), let $\mathbf{x} = (x_1, x_2, \cdots, x_k) \in X^k$ and $\mathbf{y} = (y_1, y_2, \cdots, y_k) \in Y^k$. Then the pointed Gromov–Hausdorff metric between $(X, d_X, \mathbf{x})$ and $(Y, d_Y, \mathbf{y})$ is defined to be

$$\delta_{\mathrm{pGH}}\left((X, d_X, \mathbf{x}), (Y, d_Y, \mathbf{y})\right) := \inf_{Z,\phi,\psi} \Big\{ \delta_{\mathrm{H}}\left(\phi(X), \psi(Y)\right) \vee \max_{1\le i\le k} d_Z\big(\phi(x_i), \psi(y_i)\big) \Big\},$$

where the infimum ranges over all choices of metric spaces Z and isometric embeddings $\phi : X \to Z$ and $\psi : X' \to Z$. Let $\mathbb{M}_c^k$ denote the isometry-equivalence classes of those compact metric spaces with k marked points. It is a Polish space when endowed with δ_{pGH}.

4.2.4 Real trees

A *real tree* is a geodesic metric space without loops. More precisely, a metric space (X, d, r) is called a (rooted) real tree if $r \in X$ and

- for any two points $x, y \in X$, there exists a continuous injective map $\phi_{xy} : [0, d(x, y)] \to X$ such that $\phi_{xy}(0) = x$ and $\phi_{xy}(d(x, y)) = y$. The image of ϕ_{xy} is denoted by $[\![x, y]\!]$;
- if $q : [0, 1] \to X$ is a continuous injective map such that $q(0) = x$ and $q(1) = y$, then $q([0, 1]) = [\![x, y]\!]$.

As for discrete trees, when it is clear from context which metric we are talking about, we refering to metric spaces by the sets. For instance $(\mathcal{T}, d)$ is often referred to as $\mathcal{T}$.

A *measured (rooted) real tree* is a real tree (X, d, r) equipped with a finite (Borel) measure $\mu \in \mathcal{M}(X)$. We always assume that the metric space (X, d) is complete and separable. We denote by $\mathbb{T}_w$ the set of the weak isometry equivalence classes of measured rooted real trees, equipped with the pointed Gromov–Prokhorov topology. Also, let $\mathbb{T}_w^c$ be the set of the measure-preserving isometry equivalence classes of those measured rooted real trees (X, d, r, μ) such that (X, d) is compact. We endow $\mathbb{T}_w^c$ with the pointed Gromov–Hausdorff–Prokhorov distance. Then both $\mathbb{T}_w$ and $\mathbb{T}_w^c$ are Polish spaces. However in our proofs, we do not always distinguish an equivalence class and the elements in it.

Let (T, d, r) be a rooted real tree. For $u \in T$, the degree of u in T, denoted by $\deg(u, T)$, is the number of connected components of $T \setminus \{u\}$. We also denote by

$$\mathrm{Lf}(T) = \{u \in T : \deg(u, T) = 1\} \quad \text{and} \quad \mathrm{Br}(T) = \{u \in T : \deg(u, T) \geq 3\}$$

the set of the *leaves* and the set of *branch points* of T, respectively. The skeleton of T is the complementary set of $\mathrm{Lf}(T)$ in T, denoted by $\mathrm{Sk}(T)$. For two points $u, v \in T$, we denote by $u \wedge v$ the closest common ancestor of u and v, that is, the unique point w of $[\![r, u]\!] \cap [\![r, v]\!]$ such that $d(u, v) = d(u, w) + d(w, v)$.

For a rooted real tree (T, r), if $x \in T$ then the subtree of T above x, denoted by $\mathrm{Sub}(T, x)$, is defined to be

$$\mathrm{Sub}(T, x) := \{u \in T : x \in [\![r, u]\!]\}.$$

Spanning subtree. Let (T, d, r) be a rooted real tree and let $\mathbf{x} = (x_1, \cdots, x_k)$ be k points of T for some $k \geq 1$. We denote by $\mathrm{Span}(T; \mathbf{x})$ the smallest connected set of T which contains the root r and $\mathbf{x}$, that is, $\mathrm{Span}(T; \mathbf{x}) = \cup_{1 \leq i \leq k} [\![r, x_i]\!]$. We consider $\mathrm{Span}(T; \mathbf{x})$ as a real tree rooted at r and refer to it as a *spanning subtree* or a *reduced tree* of T.

If (T, d, r) is a real tree and there exists some $\mathbf{x} = (x_1, x_2, \cdots, x_k) \in T^k$ for some $k \geq 1$ such that $T = \mathrm{Span}(T; \mathbf{x})$, then the metric aspect of T is rather simple to visualize. More precisely, if we write $x_0 = r$ and let $\rho^{\mathbf{x}} = (d(x_i, x_j), 0 \leq i, j \leq k)$, then $\rho^{\mathbf{x}}$ determines (T, d, r) under an isometry.

Gluing. If $(T_i, d_i), i = 1, 2$ are two real trees with some distinguished points $x_i \in T_i, i = 1, 2$, the result of the *gluing* of T_1 and T_2 at (x_1, x_2) is the metric space $(T_1 \cup T_2, \delta)$, where the distance δ is defined by

$$\delta(u, v) = \begin{cases} d_i(u, v), & \text{if } (u, v) \in T_i^2, i = 1, 2; \\ d_1(u, x_1) + d_2(v, x_2), & \text{if } u \in T_1, v \in T_2. \end{cases}$$

It is easy to verify that $(T_1 \cup T_2, \delta)$ is a real tree with x_1 and x_2 identified as one point, which we denote by $T_1 \circledast_{x_1 = x_2} T_2$ in the following. Moreover, if T_1 is rooted at some point r, we make the convention that $T_1 \circledast_{x_1 = x_2} T_2$ is also rooted at r.

4.2.5 Inhomogeneous continuum random trees

The inhomogeneous continuum random tree (abbreviated as ICRT in the following) has been introduced in [41] and [13]. See also [12, 15, 17] for studies of ICRT and related problems.

Let $\mathbf{\Theta}$ (the *parameter space*) be the set of sequences $\boldsymbol{\theta} = (\theta_0, \theta_1, \theta_2, \cdots) \in \mathbb{R}_+^\infty$ such that $\theta_1 \geq \theta_2 \geq \theta_3 \cdots \geq 0$, $\theta_0 \geq 0$, $\sum_{i\geq 0} \theta_i^2 = 1$, and either $\theta_0 > 0$ or $\sum_{i\geq 1} \theta_i = \infty$.

Poisson point process construction. For each $\boldsymbol{\theta} \in \mathbf{\Theta}$, we can define a real tree $\mathcal{T}$ in the following way.

- If $\theta_0 > 0$, let $\mathrm{P}_0 = \{(u_j, v_j), j \geq 1\}$ be a Poisson point process on the first octant $\{(x, y) : 0 \leq y \leq x\}$ of intensity measure $\theta_0^2 dxdy$, ordered in such a way that $u_1 < u_2 < u_3 < \cdots$.
- For every $i \geq 1$ such that $\theta_i > 0$, let $\mathrm{P}_i = \{\xi_{i,j}, j \geq 1\}$ be a homogeneous Poisson process on $\mathbb{R}_+$ of intensity θ_i under $\mathbf{P}$, such that $\xi_{i,1} < \xi_{i,2} < \xi_{i,3} < \cdots$.

All these Poisson processes are supposed to be mutually independent and defined on some common probability space $(\Omega, \mathcal{F}, \mathbf{P})$. We consider the points of all these processes as marks on the half line $\mathbb{R}_+$, among which we distinguish two kinds: the *cutpoints* and the *joinpoints*. A cutpoint is either u_j for some $j \geq 1$ or $\xi_{i,j}$ for some $i \geq 1$ and $j \geq 2$. For each cutpoint x, we associate a joinpoint x^* as follows: $x^* = v_j$ if $x = u_j$ for some $j \geq 1$ and $x^* = \xi_{i,1}$ if $x = \xi_{i,j}$ for some $i \geq 1$ and $j \geq 2$. One easily verifies that the hypotheses on $\boldsymbol{\theta}$ imply that the set of cutpoints is a.s. finite on each compact set of $\mathbb{R}_+$, while the joinpoints are dense a.s. everywhere. (See for example [13] for a proof.) In particular, we can arrange the cutpoints in increasing order as $0 < \eta_1 < \eta_2 < \eta_3 < \cdots$. This splits $\mathbb{R}_+$ into countaly intervals that we now reassemble into a tree. We write η_k^* for the joinpoint associated to the k-th cutpoint η_k. We define R_1 to be the metric space $[0, \eta_1]$ rooted at 0. For $k \geq 1$, we let

$$R_{k+1} := R_k \underset{\eta_k^* = \eta_k}{\circledast} [\eta_k, \eta_{k+1}].$$

In words, we graft the intervals $[\eta_k, \eta_{k+1}]$ by gluing the left end at the joinpoint η_k^*. Note that we have $\eta_k^* < \eta_k$ a.s., thus $\eta_k^* \in R_k$ and the above grafting operation is well defined almost surely. It follows from this Poisson construction that $(R_k)_{k\geq 1}$ is a consistent family of "discrete" trees which also verifies the "leaf-tight" condition in Aldous [10]. Therefore by [10, Theorem 3], the complete metric space $\mathcal{T} := \overline{\cup_{k\geq 1} R_k}$ is a real tree and almost surely there exists a probability measure μ, called the *mass measure*, which is concentrated on the leaf set of $\mathcal{T}$. Moreover, if conditional on $\mathcal{T}$, $(V_k, k \geq 1)$ is a sequence of i.i.d. points sampled according to μ, then for each $k \geq 1$, the spanning tree $\mathrm{Span}(T; V_1, V_2, \cdots, V_k)$ has the same unconditional distribution as R_k. The distribution of the weak isometry equivalence class of $(\mathcal{T}, \mu)$ is said to be the distribution of an *ICRT of parameter* $\boldsymbol{\theta}$, which is a probability distribution on $\mathbb{T}_w$. The push-forward of the Lebesgue measure on $\mathbb{R}_+$ defines a σ-finite measure ℓ on $\mathcal{T}$, which is concentrated on $\mathrm{Sk}(T)$ and called the *length measure* of $\mathcal{T}$. Furthermore, it is not difficult to deduce the distribution of $\ell(R_1)$ from the above construction of $\mathcal{T}$:

$$\mathbf{P}\left(\ell(R_1) > r\right) = \mathbf{P}\left(\eta_1 > r\right) = e^{-\frac{1}{2}\theta_0^2 r^2} \prod_{i\geq 1} (1 + \theta_i r) e^{-\theta_i r}, \quad r > 0. \tag{4.5}$$

In the important special case when $\boldsymbol{\theta} = (1, 0, 0, \cdots)$, the above construction coincides with the line-breaking construction of the Brownian CRT in [8, Algorithm 3], that is, $\mathcal{T}$ is the Brownian CRT. This case will be referred as the Brownian case in the sequel. We notice that whenever there is an index $i \geq 1$ such that $\theta_i > 0$, the point, denoted by β_i, which corresponds to the joinpoint $\xi_{i,1}$ is a branch point of infinite degree. According to [15, Theorem 2]), θ_i is a measurable function of $(\mathcal{T}, \beta_i)$, and we refer to it as the local time of β_i in what follows.

ICRTs as scaling limits of $\boldsymbol{p}$-trees. Let $\boldsymbol{p}_n = (p_{n1}, p_{n2}, \cdots, p_{nn})$ be a probability measure on $[n]$ such that $p_{n1} \geq p_{n2} \geq \cdots \geq p_{nn} > 0$, $n \geq 1$. Define $\sigma_n \geq 0$ by $\sigma_n^2 = \sum_{i=1}^n p_{ni}^2$ and denote by T^n the corresponding $\boldsymbol{p}_n$-tree, which we view as a metric space on $[n]$ with graph distance d_{T_n}. Suppose that the sequence $(\boldsymbol{p}_n, n \geq 1)$ verifies the following hypothesis: there exists some parameter $\boldsymbol{\theta} = (\theta_i, i \geq 0)$ such that

$$\lim_{n\to\infty} \sigma_n = 0, \qquad \text{and} \qquad \lim_{n\to\infty} \frac{p_{ni}}{\sigma_n} = \theta_i, \qquad \text{for every } i \geq 1. \tag{H}$$

Then, writing $\sigma_n T^n$ for the rescaled metric space $([n], \sigma_n d_{T^n})$, Camarri and Pitman [41] have shown that

$$(\sigma_n T^n, \boldsymbol{p}_n) \xrightarrow[d,\mathrm{GP}]{n\to\infty} (\mathcal{T}, \mu), \tag{4.6}$$

where $\to_{d,\mathrm{GP}}$ denotes the convergence in distribution with respect to the Gromov–Prokhorov topology.

4.3 Main results

4.3.1 Cutting down procedures for $\boldsymbol{p}$-trees and ICRT

Consider a $\boldsymbol{p}$-tree T. We perform a cutting procedure on T by picking each time a vertex according to the restriction of $\boldsymbol{p}$ to the remaining part; however, it is more convenient for us to retain the portion of the tree that contains a random node V sampled according to $\boldsymbol{p}$ rather than the root. We denote by $L(T)$ the number of cuts necessary until V is finally picked, and let X_i, $1 \leq i \leq L(T)$, be the sequence of nodes chosen. The following identity in distribution has been already shown in [7] in the special case of the uniform Cayley tree:

$$L(T) \stackrel{d}{=} \mathrm{Card}\{\text{vertices on the path from the root to } V\}. \tag{4.7}$$

In fact, (4.7) is an immediate consequence of the following result. In the above cutting procedure, we connect the rejected parts, which are subtrees above X_i just before the cutting, by drawing an edge between X_i and X_{i+1}, $i = 1, 2, \cdots, L(T) - 1$ (see Figure 4.1 in Section 4.4). We obtain another tree on the same vertex set, which contains a path from the first cut X_1 to the random node V that we were trying to isolate. We denote by $\mathrm{cut}(T, V)$ this tree which (partially) encodes the isolating process of V. We prove in Section 4.4 that we have

$$(\mathrm{cut}(T, V), V) \stackrel{d}{=} (T, V). \tag{4.8}$$

This identity between the pairs of trees contains a lot of information about the distributional structure of the $\boldsymbol{p}$-trees, and our aim is to obtain results similar to (4.8) for ICRTs. The method we use relies on the discrete approximation of ICRT by $\boldsymbol{p}$-trees, and a first step consists in defining the appropriate cutting procedure for ICRT.

In the case of $\boldsymbol{p}$-trees, one may pick the nodes of T in the order in which they appear in a Poisson random measure. We do not develop it here but one should keep in mind that the cutting procedure may be obtained using a Poisson point process on $\mathbb{R}_+ \times T$ with intensity measure $dt \otimes \boldsymbol{p}$. In particular, this measure has a natural counterpart in the case of ICRTs, and it is according to this measure that the points should be sampled in the continuous case.

So consider now an ICRT $\mathcal{T}$. Recall that for $\boldsymbol{\theta} \neq (1, 0, \dots)$, for each $\theta_i > 0$ with $i \geq 1$, there exists a unique point, denoted by β_i, which has infinite degree. Let $\mathcal{L}$ be the measure on $\mathcal{T}$ defined by

$$\mathcal{L}(dx) := \theta_0^2 \ell(dx) + \sum_{i\geq 1} \theta_i \delta_{\beta_i}(dx), \tag{4.9}$$

which is almost surely σ-finite (Lemma 4.22). Proving that $\mathcal{L}$ is indeed the relevant cutting measure (in a sense made precise in Proposition 4.23) is the topic of Section 4.7. Conditional on $\mathcal{T}$, let $\mathcal{P}$ be a Poisson point process on $\mathbb{R}_+ \times \mathcal{T}$ of intensity measure $dt \otimes \mathcal{L}(dx)$ and let V be a μ-point on $\mathcal{T}$. We consider the elements of $\mathcal{P}$ as the successive cuts on $\mathcal{T}$ which try to isolate the random point V. For each $t \geq 0$, define

$$\mathcal{P}_t = \{x \in \mathcal{T} : \exists\, s \leq t \text{ such that } (s,x) \in \mathcal{P}\},$$

and let $\mathcal{T}_t$ be the part of $\mathcal{T}$ still connected to V at time t, that is the collection of points $u \in \mathcal{T}$ for which the unique path in $\mathcal{T}$ from V to u does not contain any element of $\mathcal{P}_t$. Clearly, $\mathcal{T}_{t'} \subset \mathcal{T}_t$ if $t' \geq t$. We set $\mathcal{C} := \{t > 0 : \mu(\mathcal{T}_{t-}) > \mu(\mathcal{T}_t)\}$. Those are the cuts which contribute to the isolation of V.

4.3.2 Tracking one node and the one-node cut tree

We construct a tree which encodes this cutting process in a similar way that the tree $H = \mathrm{cut}(T, V)$ encodes the cutting procedure for discrete trees. First we construct the "backbone", which is the equivalent of the path we add in the discrete case. For $t \geq 0$, we define

$$L_t := \int_0^t \mu(\mathcal{T}_s)ds,$$

and L_∞ the limit as $t \to \infty$ (which might be infinite). Now consider the interval $[0, L_\infty]$, together with its Euclidean metric, that we think of as rooted at 0. Then, for each $t \in \mathcal{C}$ we graft $\mathcal{T}_{t-} \setminus \mathcal{T}_t$, the portion of the tree discarded at time t, at the point $L_t \in [0, L_\infty]$ (in the sense of the gluing introduced in Section 4.2.5). This creates a rooted real tree and we denote by $\mathrm{cut}(\mathcal{T}, V)$ its completion. Moreover, we can endow $\mathrm{cut}(\mathcal{T}, V)$ with a (possibly defective probability) measure $\hat{\mu}$ by taking the push-forward of μ under the canonical injection ϕ from $\cup_{t\in\mathcal{C}}(\mathcal{T}_{t-} \setminus \mathcal{T}_t)$ to $\mathrm{cut}(\mathcal{T}, V)$. We denote by U the endpoint L_∞ of the interval $[0, L_\infty]$. We show in Section 4.5 that

Theorem 4.4. *We have $L_\infty < \infty$ almost surely. Moreover, under (H) we have*

$$(\sigma_n \,\mathrm{cut}(T^n, V^n), \boldsymbol{p}_n, V^n) \xrightarrow[d,\mathrm{GP}]{n\to\infty} (\mathrm{cut}(\mathcal{T}, V), \hat{\mu}, U),$$

jointly with the convergence in (4.6).

Combining this with (4.8), we show in Section 4.5 that

Theorem 4.5. *Conditional on $\mathcal{T}$, U has distribution $\hat{\mu}$, and the unconditional distribution of $(\mathrm{cut}(\mathcal{T}, V), \hat{\mu})$ is the same as that of $(\mathcal{T}, \mu)$.*

Theorems 4.4 and 4.5 immediately entail that

Corollary 4.6. *Suppose that (H) holds. Then*

$$\sigma_n L(T^n) \xrightarrow[d]{n\to\infty} L_\infty,$$

jointly with the convergence in (4.6). Moreover, the unconditional distribution of L_∞ is the same as that of the distance in $\mathcal{T}$ between the root and a random point V chosen according to μ, given in (4.5).

4.3.3 The complete cutting procedure

In the procedure of the previous section, the fragmentation only takes place on the portions of the tree which contain the random point V. Following Bertoin and Miermont [30], we consider a more general cutting procedure which keeps splitting all the connected components. The aim here is to describe the genealogy of the fragmentation that this cutting procedure produces. For each $t \geq 0$, $\mathcal{P}_t$ induces an equivalence relation $\sim_t$ on $\mathcal{T}$: for $x, y \in \mathcal{T}$ we write $x \sim_t y$ if $[\![x, y]\!] \cap \mathcal{P}_t = \emptyset$. We denote by $\mathcal{T}_x(t)$ the equivalence class containing x. In particular, we have $\mathcal{T}_V(t) = \mathcal{T}_t$. Let $(V_i)_{i\geq 1}$ be a sequence of i.i.d. μ-points in $\mathcal{T}$. For each $t \geq 0$, define $\mu_i(t) = \mu(\mathcal{T}_{V_i}(t))$. We write $\boldsymbol{\mu}^{\downarrow}(t)$ for the sequence $(\mu_i(t), i \geq 1)$ rearranged in decreasing order. In the case where $\mathcal{T}$ is the Brownian CRT, the process $(\boldsymbol{\mu}^{\downarrow}(t))_{t\geq 0}$ is the fragmentation dual to the standard additive coalescent [13]. In the other cases, however, it is not even Markov because of the presence of those branch points β_i with fixed local times θ_i.

As in [30], we can define a genealogical tree for this fragmentation process. For each $i \geq 1$ and $t \geq 0$, let

$$L_t^i := \int_0^t \mu_i(s)ds,$$

and let $L_\infty^i \in [0, \infty]$ be the limit as $t \to \infty$. For each pair $(i, j) \in \mathbb{N}^2$, let $\tau(i, j) = \tau(j, i)$ be the first moment when $[\![V_i, V_j]\!]$ contains an element of $\mathcal{P}$ (or more precisely, its projection onto $\mathcal{T}$), which is almost surely finite by the properties of $\mathcal{T}$ and $\mathcal{P}$. It is not difficult to construct a sequence of increasing real trees $S_1 \subset S_2 \subset \cdots$ such that S_k has the form of a discrete tree rooted at a point denoted ρ_* with exactly k leaves $\{U_1, U_2, \cdots, U_k\}$ satisfying

$$d(\rho_*, U_i) = L_\infty^i, \quad d(U_i, U_j) = L_\infty^i + L_\infty^j - 2L_{\tau(i,j)}^i, \quad 1 \leq i < j \leq k; \tag{4.10}$$

where d denotes the distance of S_k, for each $k \geq 1$. Then we define

$$\operatorname{cut}(\mathcal{T}) := \overline{\cup_{k\geq 1} S_k},$$

the completion of the metric space $(\cup_k S_k, d)$, which is still a real tree. In the case where $\mathcal{T}$ is the Brownian CRT, the above definition of $\operatorname{cut}(\mathcal{T})$ coincides with the tree defined by Bertoin and Miermont [30].

Similarly, for each $\boldsymbol{p}_n$-tree T^n, we can define a complete cutting procedure on T^n by first generating a random permutation $(X_{n1}, X_{n2}, \cdots, X_{nn})$ on the vertex set $[n]$ and then removing X_{ni} one by one. Here the permutation $(X_{n1}, X_{n2}, \ldots, X_{nn})$ is constructed by sampling, for $i \geq 1$, X_{ni} according to the restriction of $\boldsymbol{p}_n$ to $[n] \setminus \{X_{nj}, j < i\}$. We define a new genealogy on $[n]$ by making X_{ni} an ancestor of X_{nj} if $i < j$ and X_{nj} and X_{ni} are in the same connected component when X_{ni} is removed. If we denote by $\operatorname{cut}(T^n)$ the corresponding genealogical tree, then the number of vertices in the path of $\operatorname{cut}(T^n)$ between the root X_{n1} and an arbitrary vertex v is precisely equal to the number of cuts necessary to isolate this vertex v. We have

Theorem 4.7. *Suppose that (H) holds. Then, we have*

$$\big(\sigma_n \operatorname{cut}(T^n), \boldsymbol{p}_n\big) \xrightarrow[d,\mathrm{GP}]{n\to\infty} \big(\operatorname{cut}(\mathcal{T}), \nu\big),$$

jointly with the convergence in (4.6). Here, ν is the weak limit of the empirical measures $\frac{1}{k}\sum_{i=0}^{k-1} \delta_{U_i}$, which exists almost surely conditional on $\mathcal{T}$.

From this, we show that

Theorem 4.8. *Conditionally on $\mathcal{T}$, $(U_i, i \geq 0)$ has the distribution as a sequence of i.i.d. points of common law ν. Furthermore, the unconditioned distribution of the pair $(\operatorname{cut}(\mathcal{T}), \nu)$ is the same as $(\mathcal{T}, \mu)$.*

In general, the convergence of the $\boldsymbol{p}_n$-trees to the ICRT in (4.6) cannot be improved to Gromov–Hausdorff (GH) topology, see for instance [14, Example 28]. However, when the sequence $(\boldsymbol{p}_n)_{n\geq 1}$ is suitably well-behaved, one does have this stronger convergence. (This is the case for example with $\boldsymbol{p}_n$ the uniform distribution on $[n]$, which gives rise to the Brownian CRT, see also [15, Section 4.2].) In such cases, we can reinforce accordingly the above convergences of the cut trees in the Gromov–Hausdorff topology. Note however that a "reasonable" condition on $\boldsymbol{p}$ ensuring the Gromov–Hausdorff convergence seems hard to find. Let us mention a related open question in [15, Section 7], which is to determine a practical criterion for the compactness of a general ICRT. Writing $\to_{d,\mathrm{GHP}}$ for the convergence in distribution with respect to the Gromov–Hausdorff–Prokhorov topology (see Section 4.2), we have

Theorem 4.9. *Suppose that $\mathcal{T}$ is almost surely compact and suppose also as $n \to \infty$,*

$$\big(\sigma_n T^n, \boldsymbol{p}_n\big) \underset{d,\mathrm{GHP}}{\overset{n\to\infty}{\longrightarrow}} \big(\mathcal{T}, \mu\big). \tag{4.11}$$

Then, jointly with the convergence in (4.11), we have

$$\big(\sigma_n \operatorname{cut}(T^n, V^n), \boldsymbol{p}_n\big) \underset{d,\mathrm{GHP}}{\overset{n\to\infty}{\longrightarrow}} \big(\operatorname{cut}(\mathcal{T}, V), \hat{\mu}\big),$$
$$\big(\sigma_n \operatorname{cut}(T^n), \boldsymbol{p}_n\big) \underset{d,\mathrm{GHP}}{\overset{n\to\infty}{\longrightarrow}} \big(\operatorname{cut}(\mathcal{T}), \nu\big).$$

4.3.4 Reversing the cutting procedure

We also consider the transformation that "reverses" the construction of the trees $\operatorname{cut}(\mathcal{T}, V)$ defined above. Here, by reversing we mean to obtain a tree distributed as the primal tree $\mathcal{T}$, conditioned on the cut tree being the one we need to transform. So for an ICRT $(\mathcal{H}, d_{\mathcal{H}}, \hat{\mu})$ and a random point U sampled according to its mass measure $\hat{\mu}$, we should construct a tree $\operatorname{shuff}(\mathcal{H}, U)$ such that

$$(\mathcal{T}, \operatorname{cut}(\mathcal{T}, V)) \overset{d}{=} (\operatorname{shuff}(\mathcal{H}, U), \mathcal{H}). \tag{4.12}$$

This reverse transformation is the one described in [7] for the Brownian CRT. For $\mathcal{H}$ rooted at $r(\mathcal{H})$, the path between $[\![r(\mathcal{H}), U]\!]$ that joins $r(\mathcal{H})$ to U in $\mathcal{H}$ decomposes the tree into countably many subtrees of positive mass

$$F_x = \{y \in \mathcal{H} : U \wedge y = x\},$$

where $U \wedge y$ denotes the closest common ancestor of U and y, that is the unique point a such that $[\![r(\mathcal{H}), U]\!] \cap [\![r(\mathcal{H}), y]\!] = [\![r(\mathcal{H}), a]\!]$. Informally, the tree $\operatorname{shuff}(\mathcal{H}, U)$ is the metric space one obtains from $\mathcal{H}$ by attaching each F_x of positive mass at a random point A_x, which is sampled proportionally to $\hat{\mu}$ in the union of the F_y for which $d_{\mathcal{H}}(U, y) < d_{\mathcal{H}}(U, x)$. We postpone the precise definition of $\operatorname{shuff}(\mathcal{H}, U)$ until Section 4.6.1.

The question of reversing the complete cut tree $\operatorname{cut}(\mathcal{T})$ is more delicate and is the subject of Chapter 5. There we restrict ourselves to the case of a Brownian CRT: for $\mathcal{T}$ and $\mathcal{G}$ Brownian CRT we construct a tree $\operatorname{shuff}(\mathcal{G})$ such that

$$(\mathcal{T}, \operatorname{cut}(\mathcal{T})) \overset{d}{=} (\operatorname{shuff}(\mathcal{G}), \mathcal{G}).$$

We believe that the construction there is also valid for more general ICRTs, but the arguments we use there strongly rely on the self-similarity of the Brownian CRT.

Remarks. i. Theorem 4.5 generalizes Theorem 1.5 in [7], which is about the Brownian CRT. The special case of Theorem 4.4 concerning the convergence of uniform Cayley trees to the Brownian CRT is also found there.

ii. When $\mathcal{T}$ is the Brownian CRT, Theorem 4.8 has been proven by Bertoin and Miermont [30]. Their proof relies on the self-similar property of the Aldous–Pitman's fragmentation. They also proved a convergence similar to the one in Theorem 4.7 for the conditioned Galton–Watson trees with finite-variance offspring distributions. Let us point out that their definition of the discrete cut trees is distinct from ours, and there is no "duality" at the discrete level for their definitions. Very recently, a result related to Theorem 4.7 has been proved for the case of stable trees [46] (with a different notion of discrete cut tree). Note also that the convergence of the cut trees proved in [30] and [46] is with respect to the Gromov–Prokhorov topology, so is weaker than the convergence of the corresponding conditioned Galton–Watson trees, which holds in the Gromov–Hausdorff–Prokhorov sense. In our case, the identities imply that the convergence of the cut trees is as strong as that of the $\boldsymbol{p}_n$-trees (Theorem 4.9).

iii. Abraham and Delmas [5] have shown an analog of Theorem 4.5 for the Lévy tree, introduced in [83]. In passing Aldous et al. [15] have conjured that a Lévy tree is a mixture of the ICRTs where the parameters $\boldsymbol{\theta}$ are chosen according to the distribution of the jumps in the bridge process of the associated Lévy process. Then the similarity between Theorem 4.5 and the result of Abraham and Delmas may be seen as a piece of evidence supporting this conjecture.

4.4 Cutting down and rearranging a $\boldsymbol{p}$-tree

As we have mentioned in the introduction, our approach to the theorems about continuum random trees involves taking limits in the discrete world. In this section, we prove the discrete results about the decomposition and the rearrangement of $\boldsymbol{p}$-trees that will enable us to obtain similar decomposition and rearrangement procedures for inhomogeneous continuum random trees.

4.4.1 Isolating one vertex

As a warm up, and in order to present many of the important ideas, we start by isolating a single node. Let T be a $\boldsymbol{p}$-tree and let V be an independent $\boldsymbol{p}$-node. We isolate the vertex V by removing each time a random vertex of T and preserving only the component containing V until the time when V is picked.

THE 1-CUTTING PROCEDURE AND THE 1-CUT TREE. Initially, we have $T_0 = T$, and an independent vertex V. Then, for $i \geq 1$, we choose a node X_i according to the restriction of $\boldsymbol{p}$ to the vertex set $\mathfrak{v}(T_{i-1})$ of T_{i-1}. We define T_i to be the connected component of the forest induced by T_{i-1} on $\mathfrak{v}(T_{i-1}) \setminus \{X_i\}$ which contains V. If $T_i = \varnothing$, or equivalently $X_i = V$, the process stops and we set $L = L(T) = i$. Since at least one vertex is removed at every step, the process stops in time $L \leq n$.

As we destruct the tree T to isolate V by iteratively pruning random nodes, we construct a tree which records the history of the destruction, that we call the 1-cut tree. This 1-cut tree will, in particular, give some information about the number of cuts which were needed to isolate V. However, we remind the reader that this number of cuts is not our main objective, and that we are after a more detailed correspondence between the initial tree and its 1-cut tree. We will prove that these two trees are *dual* in a sense that we will make precise shortly.

By construction, $(T_i, 0 \leq i < L)$ is a decreasing sequence of nonempty trees which all contain V, and $(X_i, 1 \leq i \leq L)$ is a sequence of distinct vertices of $T = T_0$. For $1 \leq i \leq L$, we set $F_i = T_{i-1} \setminus T_i$, that is, F_i is the graph on the vertex set $\mathfrak{v}(T_{i-1}) \setminus \mathfrak{v}(T_i)$ whose edge set is a subset of the edge set of T_{i-1}. It is not difficult to see that F_i is a tree containing X_i, which we see as the root of F_i. Besides, for each $1 \leq i < L$, $X_i \neq V$ and there is a neighbor U_i of X_i on the path between X_i and V in T_{i-1}. Then $U_i \in T_i$ and we see T_i as rooted at U_i.

When the procedure stops, we have a vector $(F_i, 1 \leq i \leq L)$ of subtrees of T which together span all of $[n]$. We may re-arrange them into a new tree, the 1-cut tree corresponding to the isolation of V in T. We do this by connecting their roots $X_1, X_2, \dots, X_L$ into a path (in this order). The resulting tree, denoted by H is seen as rooted at X_1, and carries a distinguished path or backbone $[\![X_1, V]\!]$, which we denote by S, and distinguished points $U_1, \dots, U_{L-1}$.

Note that for $i = 1, \dots, L-1$, we have $U_i \in T_i$. Equivalently, U_i lies in the subtree of H rooted at X_{i+1}. In general, for a tree $t \in \mathbb{T}_n$ and $v \in [n]$, let $x_1, \dots, x_\ell = v$ be the nodes of $\mathrm{Span}(t; v)$. We define $\mathbb{U}(t, v)$ as the collection of vectors $(u_1, \dots, u_{\ell-1})$ of nodes of $[n]$ such that $u_i \in \mathrm{Sub}(t, x_{i+1})$, for $1 \leq i < \ell$. Then by construction, for a $h \in \mathbb{T}_n$, conditional on $H = h$ and $V = v$, we have L equal to the

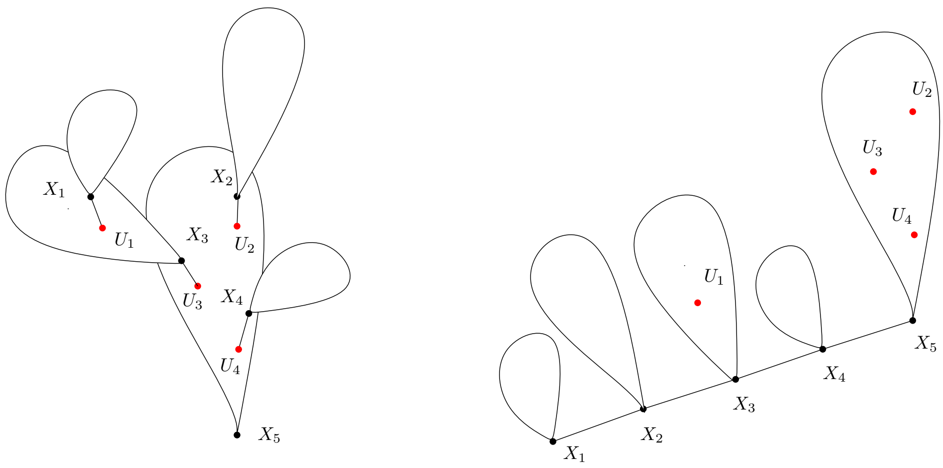

***Figure 4.1** – The re-organization of the tree in the one-cutting procedure: on the left the initial tree T, on the right H and the marked nodes $U_1, \ldots, U_4$ where to reattach $X_1, \ldots, X_4$ in order to recover T.*

number of the nodes in $\mathrm{Span}(h; v)$ and $(U_1, \ldots, U_{L-1}) \in \mathbb{U}(h, v)$ with probability one. For $A \subseteq [n]$, we write $\boldsymbol{p}(A) := \sum_{i \in A} p_i$.

Lemma 4.10. *Let T be a $\mathbf{p}$-tree on $[n]$, and V be an independent $\mathbf{p}$-node. Let $h \in \mathbb{T}_n$, and $v \in [n]$ for which $\mathrm{Span}(h; v)$ is the path made of the nodes $x_1, x_2, \ldots, x_{\ell-1}, x_\ell = v$. Let $(u_1, \ldots, u_{\ell-1}) \in \mathbb{U}(h, v)$ and $w \in [n]$. Then we have*

$$\mathbf{P}\left(H = h; V = v; r(T) = w; U_i = u_i, 1 \leq i < \ell\right) = \pi(h) \cdot \prod_{1 \leq i < \ell} \frac{p_{u_i}}{\boldsymbol{p}(\mathrm{Sub}(h, x_{i+1}))} \cdot p_v \cdot p_w.$$

In particular, $(H, V) \sim \pi \otimes \boldsymbol{p}$.

As a direct consequence of our construction of H, L is the number of nodes of the subtree $\mathrm{Span}(H, V)$, which we write $\#\,\mathrm{Span}(H, V)$. So Lemma 4.10 entails immediately that

Proposition 4.11. *Let T be a $\mathbf{p}$-tree and V be an independent $\mathbf{p}$-node. Then*

$$L \stackrel{d}{=} \#\,\mathrm{Span}(T, V).$$

Proof of Lemma 4.10. By construction, we have

$$\{H = h; V = v\} \subset \{X_1 = x_1, \cdots, X_{\ell-1} = x_{\ell-1}, X_\ell = v; L = \ell\},$$

and the sequence $(F_i, 1 \leq i \leq \ell)$ is precisely the sequence of subtrees f_i, of h rooted at x_i, $1 \leq i \leq \ell$, that are obtained when one removes the edges $\{x_i, x_{i+1}\}$, $1 \leq i < \ell$ (the edges of the subgraph $\mathrm{Span}(h; v)$). Furthermore, given that $L = \ell$ and the sequence of cut vertices $X_i = x_i$, $1 \leq i < \ell$, in order to recover the initial tree T it suffices to identify the vertices U_i, $1 \leq i < \ell$, for which there used to be an edge $\{X_i, U_i\}$ (which yields the correct adjacencies) and the root of T. Note that U_i is a node of T_i, $1 \leq i < \ell$. However, by construction, given that $H = h$ and $V = v$, the set of nodes of T_i is precisely the set of nodes of $\mathrm{Sub}(h, x_{i+1})$, the subtree of h rooted at x_{i+1}.

For $\mathbf{u} = (u_1, \cdots, u_{\ell-1}) \in \mathbb{U}(h, v)$, define $\tau(h, v; \mathbf{u})$ as the tree obtained from h by removing the edges of $\mathrm{Span}(h; v)$, and reconnecting the pieces by adding the edges $\{x_i, u_i\}$, for all the edges $\langle x_i, x_{i+1}\rangle$ in $\mathrm{Span}(h, v)$. (In particular, the number of edges is unchanged.) We regard $\tau(h, v; \mathbf{u})$ as a tree rooted at $r = x_1$, the root of h. The the tree T may be recovered by characterizing T^r, the tree T rerooted at r, and the initial root $r(T)$. We have:

$$\begin{aligned}&\{H = h; V = v; r(T) = w; U_i = u_i, 1 \le i < \ell\}\\ &\qquad\qquad\qquad\qquad = \{T^r = \tau(h, v; \mathbf{u}); r(T) = w; X_i = x_i, 1 \le i \le \ell\}.\end{aligned}$$

It follows that, for any nodes $u_1, u_2, \dots, u_{\ell-1}$ as above, we have

$$\begin{aligned}&\mathbf{P}\left(H = h; V = v; r(T) = w; U_i = u_i, 1 \le i < \ell\right)\\ &= \mathbf{P}\left(T = \tau(h, v; \mathbf{u})^w; V = v; X_i = x_i; 1 \le i \le \ell\right)\\ &= \pi(\tau(h, v; \mathbf{u})^w) \cdot p_v \cdot \prod_{1 \le i \le \ell} \frac{p_{x_i}}{\boldsymbol{p}(\mathrm{Sub}(h, x_i))}.\end{aligned}$$

Now, by definition, the only nodes that get their (in-)degree modified in the transformation from h to $\tau(h, v; \mathbf{u})$ are u_i, x_{i+1}, $1 \le i < \ell$: every such x_{i+1} gets one less in-edge while u_i gets one more. The re-rooting at w then only modifies the in-degrees of the extremities of the path that is reversed, namely $x_1 = r$ and w. It follows that

$$\pi(\tau(h, v; \mathbf{u})^w) = \pi(h) \cdot \prod_{1 \le i < \ell} \frac{p_{u_i}}{p_{x_{i+1}}} \cdot \frac{p_w}{p_{x_1}}.$$

Since $\boldsymbol{p}(\mathrm{Sub}(h, x_1)) = 1$, we have

$$\mathbf{P}\left(H = h; V = v; r(T) = w; U_i = u_i, 1 \le i < \ell\right) = \pi(h) \cdot \prod_{1 \le i < \ell} \frac{p_{u_i}}{\boldsymbol{p}(\mathrm{Sub}(h, x_{i+1}))} \cdot p_v \cdot p_w,$$

which proves the first claim. Summing over all the choices for $\mathbf{u} = (u_1, u_2, \dots, u_{\ell-1}) \in \mathbb{U}(h, v)$, and $w \in [n]$, we obtain

$$\begin{aligned}\mathbf{P}\left(H = h; V = v\right) &= \sum_{w \in [n]} \sum_{\mathbf{u} \in \mathbb{U}(h,v)} \pi(h) \cdot \prod_{1 \le i < \ell} \frac{p_{u_i}}{\boldsymbol{p}(\mathrm{Sub}(h, x_{i+1}))} \cdot p_v \cdot p_w\\ &= \pi(h) \cdot p_v \cdot \sum_{\substack{\mathbf{u} = (u_1, \cdots, u_{\ell-1}):\\ u_i \in \mathrm{Sub}(h, x_{i+1}), 1 \le i < \ell}} \frac{p_{u_1}}{\boldsymbol{p}(\mathrm{Sub}(h, x_2))} \cdots \frac{p_{u_{\ell-1}}}{\boldsymbol{p}(\mathrm{Sub}(h, x_\ell))}\\ &= \pi(h) \cdot p_v,\end{aligned}$$

which completes the proof. □

THE REVERSE 1-CUTTING PROCEDURE. We have transformed the tree T into the tree H, by somewhat "knitting" a path between the first picked random $\boldsymbol{p}$-node X_1 and the distinguished node V. This transform is reversible. Indeed, it is possible to "unknit" the path between V and the root of H, and reshuffle the subtrees thereby created in order to obtain a new tree $\tilde{T}$, distributed as T and in which V is an independent $\boldsymbol{p}$-node. Knowing the U_i, one could do this exactly, and recover the adjacencies of T (recovering T also requires the information about the root $r(T)$ which has been lost). Defining a reverse transformation reduces to finding the joint distribution of (U_i) and $r(T)$, which is precisely the statement of Lemma 4.10, so that the following reverse construction is now straightforward.

Let $h \in \mathbb{T}_n$, rooted at r and let v be a node in $[n]$. We think of h as the tree that was obtained by the 1-cutting procedure $\mathrm{cut}(T, v)$, for some initial tree T. Suppose that $\mathrm{Span}(h, v)$ consists of the vertices $r = x_1, x_2, \dots, x_\ell = v$. Removing the edges of $\mathrm{Span}(h, v)$ from h disconnects it into ℓ connected components which we see as rooted at x_i, $1 \le i \le \ell$. For $w \in \mathrm{Span}^*(h, v) = \mathrm{Span}(h, v) \setminus \{r\}$, sample a node U_w according to the restriction of $\boldsymbol{p}$ to $\mathrm{Sub}(h,w)$. Let $\mathbf{U} = (U_w, w \in \mathrm{Span}^*(h,v))$ be the obtained vector. Then $\mathbf{U} \in \mathbb{U}(h,v)$. We then define $\mathrm{shuff}(h,v)$ to be the rooted tree which has the adjacencies of $\tau(h,v;\mathbf{U})$, but that is re-rooted at an independent $\boldsymbol{p}$-node.

It should now be clear that the 1-cutting procedure and the reshuffling operation we have just defined are dual in the following sense.

Proposition 4.12 (1-cutting duality). *Let T be $\boldsymbol{p}$-tree on $[n]$ and V be an independent $\boldsymbol{p}$-node. Then,*

$$(\mathrm{shuff}(T,V), T, V) \stackrel{d}{=} (T, \mathrm{cut}(T,V), V).$$

In particular, $(\mathrm{shuff}(T,V), V) \sim \pi \otimes \boldsymbol{p}$.

Note that for the joint distribution in Proposition 4.12, it is necessary to re-root at another independent $\boldsymbol{p}$-node in order to have the claimed equality. Indeed, T and $\tau(T,V;\mathbf{U})$ have the same root almost surely, while T and $\mathrm{cut}(T,V)$ do not (they only have the same root with probability $\sum_{i\geq 1} p_i^2 < 1$).

Proof of Proposition 4.12. Let $H = \mathrm{cut}(T,V)$ be the tree resulting from the cutting procedure. Let $L = \#\mathrm{Span}(H;V)$. For $1 \leq i < L$, we defined nodes U_i, which used to be the neighbors of X_i in T. For $w \in \mathrm{Span}^*(H;V)$, we let $U_w = U_i$ if $w = X_{i+1}$, and let $\mathbf{U}$ be the corresponding vector. Then writing $\hat{r} = r(T)$, with probability one, we have

$$T = \tau(H,V;\mathbf{U})^{\hat{r}}.$$

By Lemma 4.10, $\mathbf{U} \in \mathbb{U}(H,V)$ and conditional on H and V, U_w, $w \in \mathrm{Span}^*(H,V)$ and $\hat{r} = r(T)$ are independent and distributed according to the restriction of $\boldsymbol{p}$ to $\mathrm{Sub}(H,w)$ and $\boldsymbol{p}$, respectively. So this coupling indeed gives that $T = \tau(H,V;\mathbf{U})^{\hat{r}}$ is distributed as $\mathrm{shuff}(H,V)$, conditional on H. Since in this coupling $(\mathrm{shuff}(H,V), T, V)$ is almost surely equal to (T, H, V), the proof is complete. $\square$

Remark. Note that the shuffle procedure would permit to obtain the original tree T exactly if we were to use some information that might be gathered as the cutting procedure goes on. In this discrete case, this is rather clear that one could do this, since the shuffle construction only consists in replacing some edges with others but the vertex set remains the same. This observation will be used in Section 4.6 to prove a similar statement for the ICRT. There it is much less clear and the result is slightly weaker: it is possible to couple the shuffle in such a way that the tree obtained is measure-isometric to the original one.

4.4.2 Isolating multiple vertices

We define a cutting procedure analogous to the one described in Section 4.4.1, but which continues until multiple nodes have been isolated. Again, we let T be a $\boldsymbol{p}$-tree and, for some $k \geq 1$, let $V_1, V_2, \cdots, V_k$ be k independent vertices chosen according to $\boldsymbol{p}$ (so not necessarily distinct).

THE k-CUTTING PROCEDURE AND THE k-CUT TREE. We start with $\Gamma_0 = T$. Later on, Γ_i is meant to be the forest induced by T on the nodes that are left. For each time $i \geq 1$, we pick a random vertex X_i according to $\boldsymbol{p}$ restricted to $\mathfrak{v}(\Gamma_{i-1})$, the set of the remaining vertices, and remove it. Then among the connected components of $T \setminus \{X_1, \cdots, X_i\}$, we only keep those containing at least one of $V_1, \cdots, V_k$. We stop at the first time when all k vertices $V_1, \ldots, V_k$ have been chosen, that is at time

$$L^k := \inf\{i \geq 1 : \{V_1, \ldots, V_k\} \subseteq \{X_1, \ldots, X_i\}\}.$$

For $1 \leq \ell \leq k$ and for $i \geq 0$, we denote by T_i^ℓ the connected component of $T \setminus \{X_1, X_2, \cdots, X_i\}$ containing V_ℓ at time i, or $T_i^\ell = \varnothing$ if $V_\ell \in \{X_1, \ldots, X_i\}$. Then Γ_i is the graph consisting of the connected components T_i^ℓ, $\ell = 1, \ldots, k$.

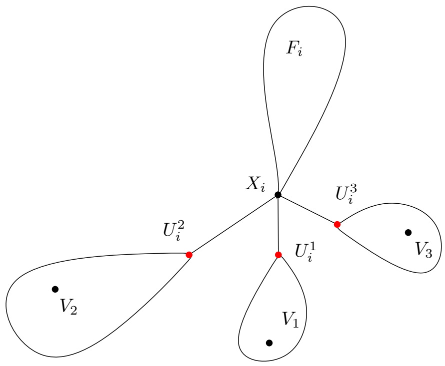

Figure 4.2 – *The decomposition of the tree when removing the point X_i from the connected component of Γ_i which contains V_1, V_2 and V_3.*

Fix some $\ell \in \{1, 2, \dots, k\}$, and suppose that at time $i \geq 1$, we have $X_i \in T_{i-1}^\ell$. If $X_i = V_\ell$, then $T_i^\ell = \varnothing$ and we define $F_i = T_{i-1}^\ell$, re-rooted at $X_i = V_\ell$. Otherwise, $X_i \neq V_\ell$ and there is a first node U_i^ℓ on the path between X_i and V_ℓ in T_{i-1}^ℓ. Then $U_i^\ell \in T_i^\ell$, and we see T_i^ℓ as rooted at U_i^ℓ. Note that it is possible that $T_{i-1}^j = T_{i-1}^\ell$, for $j \neq \ell$, and that removing X_i may separate V_ℓ from V_j. Removing from Γ_{i-1} the edges $\{X_i, U_i^\ell\}$, for $1 \leq \ell \leq k$ such that $T_i^\ell \ni X_i$, isolates X_i from the nodes $V_1, \dots, V_k$, and we define F_i as the subtree of T induced on the nodes in $\Gamma_{i-1} \setminus \Gamma_i$, so that F_i is the portion of the forest Γ_{i-1} which gets discarded at time i, which we see as rooted at X_i.

Consider the set of effective cuts which affect the size of T_i^ℓ:

$$\mathcal{E}_\ell^k = \{x \in [n] : \text{there exists } i \geq 1, \text{ such that } X_i = x \in T_{i-1}^\ell\},$$

and note that $\mathcal{E}_1^k \cup \mathcal{E}_2^k \cup \cdots \cup \mathcal{E}_k^k = \{X_i : 1 \leq i \leq L^k\}$. Let S_k, the *k-cutting skeleton*, be a tree on $\mathcal{E}_1^k \cup \cdots \cup \mathcal{E}_k^k$ that is rooted at X_1, and such that the vertices on the path from X_1 to V_ℓ in S_k are precisely the nodes of $\mathcal{E}_\ell^k$, in the order given by the indices of the cuts. So if we view S_k as a genealogical tree, then in particular, for $1 \leq j, \ell \leq k$, the common ancestors of V_j and V_ℓ are exactly the ones in $\mathcal{E}_j^k \cap \mathcal{E}_\ell^k$. The tree S_k constitutes the *backbone* of a tree on $[n]$ which we now define. For every $x \in S_k$, there is a unique $i = i(x) \geq 1$ such that $x = X_i$. For that integer i we have defined a subtree F_i which contains $X_i = x$. We append F_i to S_k at x. Formally, we consider the tree on $[n]$ whose edge set consists of the edges of S_k together with the edges of all F_i, $1 \leq i \leq L^k$. Furthermore, the tree is considered as rooted at X_1. Then this tree is completely determined by T, $V_1, \dots, V_k$, and the sequence $\mathbf{X} := (X_i, i = 1, \dots, L^k)$, and we denote this tree by $\kappa(T; V_1, \dots, V_k; \mathbf{X})$ when we want to emphasize the dependence in $\mathbf{X}$, or more simply $\mathrm{cut}(T, V_1, \dots, V_k)$ (in which it is implicit that the cutting sequence used in the transformation is such that for every $i \geq 1$, X_i is a $\boldsymbol{p}$-node in Γ_{i-1}). Clearly, if $H_k = \mathrm{cut}(T, V_1, \dots, V_k)$ then $S_k = \mathrm{Span}(H_k; V_1, \dots, V_k)$.

It is convenient to define a *canonical (total) order* $\preceq$ on the vertices of S_k. It will be needed later on in order to define the reverse procedure. For two nodes u, v in S_k, we say that $u \preceq v$ if either $u \in [\![X_1, v]\!]$, or if there exists $\ell \in \{1, \dots, k\}$ such that $u \in \mathrm{Span}(S_k; V_1, \dots, V_\ell)$ but $v \notin \mathrm{Span}(S_k; V_1, \dots, V_\ell)$.

A USEFUL COUPLING. It is useful to see all the trees $\mathrm{cut}(T; V_1, \dots, V_k)$ on the same probability space, and provide a natural but crucial coupling for which the sequence (S_k) is increasing in k. Let Y_i, $i \geq 1$, be a sequence of i.i.d. $\boldsymbol{p}$-nodes. For $k \geq 1$, we define an increasing sequence σ_k as follows. Let $\sigma_k(1) = 1$. Suppose that we have already defined $X_1^k, \dots, X_{i-1}^k$. Let Γ_{i-1}^k be the collection of connected components of $T \setminus \{X_1^k, \dots, X_{i-1}^k\}$ which contain at least one of $V_1, \dots, V_k$. Let

$$\sigma_k(i) = \inf\{j > \sigma_k(i-1) : Y_j \in \Gamma_{i-1}^k\},$$

and define $X_i^k = Y_{\sigma_k(i)}$. Then, for every k, X_i^k, $i \geq 1$, is a sequence of nodes sampled according to the restriction of $\boldsymbol{p}$ to Γ_{i-1}^k, so that $\mathbf{X}^k := (X_i^k, i \geq 1)$ can be used to define $\mathrm{cut}(T, V_1, \dots, V_k)$, $k \geq 1$, in a consistent way by setting

$$\mathrm{cut}(T, V_1, \dots, V_k) = \kappa(T, V_1, \dots, V_k; \mathbf{X}^k).$$

Suppose that the trees $H_k := \mathrm{cut}(T; V_1, \dots, V_k)$, $k \geq 1$, are constructed using the coupling we have just described. By convention let $H_0 = T$ and $\mathrm{Span}(T; \varnothing) = \varnothing$.

Lemma 4.13. *Let $S_k = \mathrm{Span}(H_k; V_1, \dots, V_k)$. Then, $S_k \subseteq S_{k+1}$ and*

$$S_k = \mathrm{Span}(S_{k+1}; V_1, \dots, V_k).$$

Proof. Let T_i^ℓ be the connected component of Γ_i^k which contains V_ℓ. Let $\hat{T}_j^\ell$ be the connected component of $T \setminus \{Y_1, Y_2, \dots, Y_j\}$ which contains V_ℓ. Then, for $\ell \leq k$, we have

$$\mathcal{E}_\ell^k = \{x : \exists i \geq 1, x = X_i^k \in T_{i-1}^\ell\} = \{y : \exists j \geq 1, y = Y_j \in \hat{T}_{j-1}^\ell\},$$

so that $\mathcal{E}_\ell^k$ does not depend on k. Then S_k is the tree on $\mathcal{E}_1^k \cup \dots \cup \mathcal{E}_k^k$ such that the nodes on the path $\mathrm{Span}(S_k; V_\ell)$ are precisely the nodes of $\mathcal{E}_\ell^k$, in the order given by the cut sequence $\mathbf{X}^k$. It follows that $S_k \subseteq S_{k+1}$ and more precisely that $S_k = \mathrm{Span}(S_{k+1}; V_1, \dots, V_k)$. $\square$

Remark. The coupling we have just defined justifies an *ordered cutting procedure* which is very similar to the one defined in [7]. Suppose that, for some $j, \ell \in \{1, \dots, k\}$ we have $x \in \mathcal{E}_j^k \setminus \mathcal{E}_\ell^k$ and $y \in \mathcal{E}_\ell^k \setminus \mathcal{E}_j^k$. Write $(\tilde{X}_i, i \geq 1)$ for the sequence in which we have exchanged the positions of x and y. Then the trees T_i^k, $i \geq \max\{m : X_m = x \text{ or } y\}$ are unaffected if we replace $(X_i, i \geq 1)$ by $(\tilde{X}_i, i \geq 1)$ in the cutting procedure. In particular, if we are only interested in the final tree H_k, we can always suppose that there exist numbers $0 = m_0 < m_1 < m_2 < \dots < m_k \leq n$ such that, for $1 \leq \ell \leq k$, and if $V_\ell \notin \{V_1, \dots, V_j\}$, we have

$$\mathcal{E}_\ell^k \setminus \bigcup_{1 \leq j < \ell} \mathcal{E}_j^k = \{X_i : m_{\ell-1} < i \leq m_\ell\}.$$

However, we prefer the coupling over the reordering of the sequence since it does not involve any modification of the distribution of the cutting sequences.

Let $\tilde{T}_k$ be the subtree of $H_{k-1} \setminus \mathrm{Span}(H_{k-1}; V_1, \dots, V_{k-1}) = H_{k-1} \setminus S_{k-1}$ which contains V_k; we agree that $\tilde{T}_k = \varnothing$ if $V_k \in \mathrm{Span}(H_{k-1}; V_1, \dots, V_{k-1})$.

Lemma 4.14. *Let T be a $\boldsymbol{p}$-tree and let V_k, $k \geq 1$, be a sequence of i.i.d. $\boldsymbol{p}$-nodes. Then, for each $k \geq 1$:*

i. Let $\mathbf{V} \subseteq [n]$ with $\mathbf{V} \neq \varnothing$, then conditional on $V_\ell \in \mathfrak{v}(\tilde{T}_k) = \mathbf{V}$, the pair $(\tilde{T}_k, V_\ell)$ is distributed as $\pi|_{\mathbf{V}} \otimes \boldsymbol{p}|_{\mathbf{V}}$, and is independent of $(H_{k-1} \setminus \mathbf{V}, V_1, \cdots, V_{k-1})$.

ii. The joint distribution of $(H_k, V_1, \cdots, V_k)$ is given by $\pi \otimes \boldsymbol{p}^{\otimes k}$.

Proof. We proceed by induction on $k \geq 1$. Let $\tilde{R}_k$ denote the tree induced by H_k on the vertex set $[n] \setminus \mathfrak{v}(\tilde{T}_k)$. For the base case $k = 1$, the first claim is trivial since $\tilde{T}_1 = T$, and the second is exactly the statement of Lemma 4.10.

Given the two subtrees $\tilde{T}_k$ and $\tilde{R}_k$, it suffices to identify where the tree $\tilde{T}_k$ is grafted on $\tilde{R}_k$ in order to recover the tree H_{k-1}. By construction, the edge connecting $\tilde{T}_k$ and $\tilde{R}_k$ in H_{k-1} binds the root of $\tilde{T}_k$ to a node of $\mathrm{Span}(\tilde{R}_k; V_1, \dots, V_{k-1})$. Let $t \in \mathbb{T}_{\mathbf{V}}$, $r \in \mathbb{T}_{[n] \setminus \mathbf{V}}$, $v_k \in \mathbf{V}$ and $v_i \in [n] \setminus \mathbf{V}$ for $1 \leq i < k$. Write $\mathbf{v}_{k-1} = \{v_1, \dots, v_{k-1}\}$. For a given node $x \in \mathrm{Span}(r; \mathbf{v}_{k-1})$, let $j_x(r, t)$ (the joint of r and t at x) be the tree obtained from t and r by adding an edge between x and the root of t. By the induction

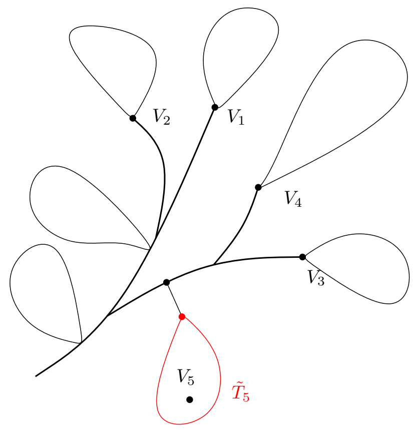

Figure 4.3 – *In order to obtain* $\mathrm{cut}(T, V_1, \dots, V_k)$ *from* $\mathrm{cut}(T, V_1, \dots, V_{k-1})$*, it suffices to transform the subtree* $\tilde{T}_k$ *of* $\mathrm{cut}(T, V_1, \dots, V_{k-1}) \setminus S_{k-1}$ *which contains* V_k*.*

hypothesis, $(H_{k-1}, V_1, \cdots, V_{k-1})$ is distributed like a $\boldsymbol{p}$-tree together with $k-1$ independent $\boldsymbol{p}$-nodes. Furthermore V_k is independent of $(H_{k-1}, V_1, \cdots, V_{k-1})$. It follows that

$$
\begin{aligned}
&\mathbb{P}(\tilde{T}_k = t; \tilde{R}_k = r; V_i = v_i, 1 \le i \le k) \\
&= \sum_{x \in \mathrm{Span}(r; \mathbf{v}_{k-1})} \mathbb{P}(H_{k-1} = j_x(r,t); V_i = v_i, 1 \le i \le k) \\
&= \sum_{x \in \mathrm{Span}(r; \mathbf{v}_{k-1})} \prod_{i \in \mathbf{V}} p_i^{C_i(t)} \cdot \prod_{j \in [n] \setminus \mathbf{V}} p_j^{C_j(r)} \cdot p_x \cdot \prod_{1 \le i \le k} p_{v_i} \\
&= \prod_{i \in \mathbf{V}} p_i^{C_i(t)} \cdot p_{v_k} \cdot \prod_{j \in [n] \setminus \mathbf{V}} p_j^{C_j(r)} \cdot \boldsymbol{p}(\mathrm{Span}(r; \mathbf{v}_{k-1})) \cdot \prod_{1 \le i < k} p_{v_i}.
\end{aligned}
$$

By summing over t and r and applying Cayley's multinomial formula, we deduce that conditional on $v(\tilde{T}_k) = \mathbf{V} \neq \varnothing$, $(\tilde{T}_k, V_k)$ is independent of $(\tilde{R}_k, V_1, \dots, V_{k-1})$ and distributed according to $\pi|_{\mathbf{V}} \otimes \boldsymbol{p}|_{\mathbf{V}}$, which establishes the first claim for k.

Now, conditional on the event $\{V_k \in S_{k-1}\}$, the vertex V_k is distributed according to the restriction of $\boldsymbol{p}$ to S_{k-1}. In this case, $H_k = H_{k-1}$ so that by the induction hypothesis

$$
\text{on } \{V_k \in S_{k-1}\}, \qquad (H_k, V_1, \dots, V_k) \sim \pi \otimes \boldsymbol{p}^{k-1} \otimes \boldsymbol{p}|_{S_{k-1}}. \tag{4.13}
$$

On the other hand, if $V_k \notin S_{k-1}$, then $\mathfrak{v}(\tilde{T}_k) \neq \varnothing$ and conditional on $\mathfrak{v}(\tilde{T}_k) = \mathbf{V}$, we have $(\tilde{T}_k, V_k) \sim \pi|_{\mathbf{V}} \otimes \boldsymbol{p}|_{\mathbf{V}}$. In that case, H_k is obtained from H_{k-1} by replacing $\tilde{T}_k$ by $\mathrm{cut}(\tilde{T}_k, V_k)$. We have already proved that, in this case, $(\tilde{T}_k, V_k)$ is independent of $\tilde{R}_k$, and Lemma 4.10 ensures that the replacement does not alter the distribution. In other words,

$$
\text{on } \{V_k \notin S_{k-1}\}, \qquad (H_k, V_1, \dots, V_k) \sim \pi \otimes \boldsymbol{p}^{k-1} \otimes \boldsymbol{p}|_{[n] \setminus S_{k-1}}. \tag{4.14}
$$

Since $V_k \sim \boldsymbol{p}$ is independent of everything else, conditional on S_{k-1}, the event $\{V_k \in S_{k-1}\}$ occurs precisely with probability $\boldsymbol{p}(S_{k-1})$, so that putting (4.13) and (4.14) together completes the proof of the induction step. $\square$

Corollary 4.15. *Suppose that* T *is a* $\boldsymbol{p}$*-tree and that* $V_1, \dots, V_k$ *are* $k \ge 1$ *independent* $\boldsymbol{p}$*-nodes, also independent of* T*. Then,*

$$
S_k \overset{d}{=} \mathrm{Span}(T; V_1, \dots, V_k).
$$

In particular, the total number of cuts needed to isolate $V_1, \dots, V_k$ *in* T *is distributed as the number of vertices of* $\mathrm{Span}(T; V_1, \dots, V_k)$*.*

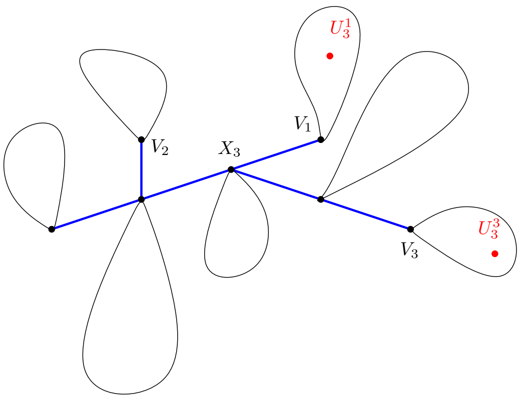

***Figure 4.4** – The 3-cut tree and the marked points U_3^1, U_3^3 corresponding to the cut node X_3. The backbone is represented by the subtree in thick blue.*

REVERSE k-CUTTING AND DUALITY. As when we were isolating a single node V in Section 4.4.1, the transformation that yields $H_k = \mathrm{cut}(T, V_1, \ldots, V_k)$ is reversible. To reverse the 1-cutting procedure, we "unknitted" the path between X_1 and V. Similarly, to reverse the k-cutting procedure, we "unknit" the backbone S_k and by doing this obtain a collection of subtrees; then we re-attach these pendant subtrees at random nodes, which are chosen in suitable subtrees in order to obtain a tree distributed like the initial tree T.

For every i, the subtree F_i, rooted at X_i, was initially attached to the set of nodes

$$\mathcal{U}_i := \{U_i^j : 1 \leq j \leq k \text{ such that } T_{i-1}^j \ni X_i\}.$$

The corresponding edges have been replaced by some edges which now lie in the backbone S_k. So, to reverse the cutting procedure knowing the sets $\mathcal{U}_i$, it suffices to remove all the edges of S_k, and to re-attach X_i to every node in $\mathcal{U}_i$. In other words, defining a reverse k-cutting transformation knowing only the tree H_k and the distinguished nodes $V_1, \ldots, V_k$ reduces to characterizing the distribution of the sets $\mathcal{U}_i$.

Consider a tree $h \in \mathbb{T}_n$, and k nodes $v_1, v_2, \ldots, v_k$ not necessarily distinct. Removing the edges of $\mathrm{Span}(h; v_1, \ldots, v_k)$ from h disconnects it into connected components f_x, each containing a single vertex x of $\mathrm{Span}(h; v_1, \ldots, v_k)$. For a given edge $\langle x, w \rangle$ of $\mathrm{Span}(h; v_1, \ldots, v_k)$, let u_w be a node in $\mathrm{Sub}(h, w)$. Let $\mathbf{u}$ be the vector of the u_w, sorted according to the canonical order of w on $\mathrm{Span}(h; v_1, \ldots, v_k)$ (see p. 107). For a given tree h and $v_1, \ldots, v_k$, we let $\mathbb{U}(h, v_1, \ldots, v_k)$ be the set of such vectors $\mathbf{u}$. For $\mathbf{u} \in \mathbb{U}(h, v_1, \ldots, v_k)$, define $\tau(h, v_1, \ldots, v_k; \mathbf{u})$ as the graph obtained from h by removing every edge $\langle x, w \rangle$ of $\mathrm{Span}(h; v_1, \ldots, v_k)$ and replacing it by $\{x, u_w\}$. We regard $\tau(h, v_1, \ldots, v_k; \mathbf{u})$ as rooted at the root of h.

Lemma 4.16. *Suppose that $h \in \mathbb{T}_n$, and that $v_1, v_2, \ldots, v_k$ are k nodes of $[n]$, not necessarily distinct. Then for every $\mathbf{u} \in \mathbb{U}(h, v_1, \ldots, v_k)$, $\tau(h, v_1, \ldots, v_k; \mathbf{u})$ is a tree on $[n]$.*

Proof. Write $t := \tau(h, v_1, \ldots, v_k; \mathbf{u})$. We proceed by induction on $n \geq 1$. For $n = 1$, $t = h$ is reduced to a single node; so t is a tree.

Suppose now that for any tree t' of size at most $n-1$, any $k \geq 1$, any nodes $v_1, v_2, \ldots, v_k \in \mathfrak{v}(t')$, and any $\mathbf{u}' \in \mathbb{U}(t', v_1, \ldots, v_k)$, the graph $\tau(t', v_1, \ldots, v_k; \mathbf{u}')$ is a tree. Let N be the set of neighbors of the root x_1 of h. For $y \in N$, define $\mathbf{v}_y$ the subset of $\{v_1, \ldots, v_k\}$ containing the vertices which lie in $\mathrm{Sub}(h, y)$. If $\mathbf{v}_y \neq \varnothing$, let also $\mathbf{u}_y \in \mathbb{U}(\mathrm{Sub}(h, y), \mathbf{v}_y)$ be obtained from $\mathbf{u}$ by keeping only the vertices u_w for $w \in \mathrm{Span}^*(\mathrm{Sub}(h, y), \mathbf{v}_y)$, still in the canonical order. Then, by construction, the subtrees $\mathrm{Sub}(h, y)$, with $y \in N$ such that $\mathbf{v}_y \neq \varnothing$ are transformed regardless of one another, and the others, for which $\mathbf{v}_y = \varnothing$, are left untouched. So the graph $\tau(h, v_1, \dots, v_k; \mathbf{u})$ induced on $[n] \setminus \{x_1\}$ consists precisely of $\tau(\mathrm{Sub}(h, y), \mathbf{v}_y; \mathbf{u}_y)$, $y \in N$. By the induction hypothesis, these subgraphs are actually trees. Then $\tau(h, v_1, \dots, v_k; \mathbf{u})$ is simply obtained by adding the node x_1 together with the edges $\{x_1, u_y\}$, for $y \in N$, where $u_y \in \mathrm{Sub}(h, y)$. In other words, each such edge connects x_1 to a different tree $\tau(\mathrm{Sub}(h, y), \mathbf{v}_y; \mathbf{u}_y)$ so that the resulting graph is also a tree. □

For a given tree h and $v_1, \dots, v_k \in [n]$ let $\mathbf{U} \in \mathbb{U}(h, v_1, \dots, v_k)$ be obtained by sampling U_w according to the restriction of $\boldsymbol{p}$ to $\mathrm{Sub}(h, w)$, for every $w \in \mathrm{Span}^*(h, v_1, \dots, v_k)$. Finally, we define the k-shuffled tree $\mathrm{shuff}(h; v_1, \dots, v_k)$ to be the tree $\tau(h, v_1, \dots, v_k; \mathbf{U})$ re-rooted at an independent $\boldsymbol{p}$-node.

We have the following result, which expresses the fact that the k-cutting and k-shuffling procedures are truly reverses of one another.

Proposition 4.17 (k-cutting duality). *Let T be a $\mathbf{p}$-tree and let $V_1, \dots, V_k$ be k independent $\mathbf{p}$-nodes, also independent of T. Then, we have the following duality*

$$(\mathrm{shuff}(T, V_1, \dots, V_k), T, V_1, \dots, V_k) \stackrel{d}{=} (T, \mathrm{cut}(T, V_1, \dots, V_k), V_1, \dots, V_k).$$

In particular, $(\mathrm{shuff}(T, V_1, \dots, V_k), V_1, \dots, V_k) \sim \pi \otimes \boldsymbol{p}^{\otimes k}$.

Proof. We consider the coupling we have defined on page 107: We let $H_k = \mathrm{cut}(T, V_1, \dots, V_k)$ for a $\boldsymbol{p}$-tree T rooted at $\hat{r} = r(T)$, and for every edge $\langle x, w \rangle$ of $\mathrm{Span}(H_k; V_1, \dots, V_k)$ we let U_w be the unique node of $\mathrm{Sub}(H_k, w)$ which used to be connected to x in the initial tree T. This defines the vector $\mathbf{U} = (U_w, w \in \mathrm{Span}^*(H_k; V_1, \dots, V_k))$. We show by induction on $k \geq 1$ that $\tau(H_k, V_1, \dots, V_k; \mathbf{U})^{\hat{r}} = T$ and that the joint distribution of $(H_k, \hat{r}, V_1, \dots, V_k, \mathbf{U})$ is that required by the construction above, so that

$$(\tau(H_k, V_1, \dots, V_k; \mathbf{U})^{\hat{r}}, H_k, V_1, \dots, V_k) \stackrel{d}{=} (\mathrm{shuff}(H_k, V_1, \dots, V_k), H_k, V_1, \dots, V_k).$$

Since $H_k \stackrel{d}{=} T$, this would complete the proof.

For $k = 1$, the statement corresponds precisely to the construction of the proof of Proposition 4.12. As before, for $\ell \leq k$, we let $S_\ell = \mathrm{Span}(H_k; V_1, \dots, V_\ell)$. If $k \geq 2$, let $\tilde{R}_k$ be the connected component of $H_k \setminus S_{k-1}$ which contains V_k, or $\tilde{R}_k = \varnothing$ if $V_k \in S_{k-1}$. In the latter case, $T = \tau(H_k, V_1, \dots, V_{k-1}, \mathbf{U})^{\hat{r}}$ and the joint distribution of $(H_k, \hat{r}, V_1, \dots, V_{k-1}, \mathbf{U})$ is correct by the induction hypothesis. Otherwise, let $\mathbf{U}_k$ denote the sub-vector of $\mathbf{U}$ consisting of the components U_w for $w \in \mathrm{Span}^*(\tilde{R}_k, V_k)$, and let $\mathbf{U}_{1,k-1} = (U_w, w \in \mathrm{Span}^*(H_k; V_1, \dots, V_{k-1}))$. If $\theta \in S_k$ is the unique point such that $\tilde{R}_k = \mathrm{Sub}(H_k, \theta)$ (that is, θ is the root of $\tilde{R}_k$), then removing $\tilde{R}_k$ from H_k and replacing it by $\tau(\tilde{R}_k, V_k; \mathbf{U}_k)^{U_\theta}$ yields precisely the tree $H_{k-1} := \mathrm{cut}(T; V_1, \dots, V_{k-1})$. Also, the distribution of $(\tilde{R}_k, U_\theta, V_k, \mathbf{U}_k)$ is correct, since conditional on the vertex set $\tilde{R}_k$ is distributed as $\pi|_{\mathfrak{v}(\tilde{R}_k)}$ (see (i) of Lemma 4.14). Note that this transformation does not modify the distribution of $\mathbf{U}_{1,k-1}$. By the induction hypothesis, $T = \tau(H_{k-1}, V_1, \dots, V_{k-1}; \mathbf{U}_{1,k-1})^{\hat{r}}$. Since conditionally on $S_{k-1} = \mathrm{Span}(H_k; V_1, \dots, V_{k-1})$ we have $V_k \in S_{k-1}$ with probability $\boldsymbol{p}(S_{k-1})$, the proof is complete. □

4.4.3 The complete cutting and the cut tree.

For n a natural number, we may also easily apply the previous procedure until all n nodes have been chosen. In this case, the cutting procedure continues recursively in *all* the connected components. The number of cuts is now completely irrelevant (it is a.s. equal to n), and we define the forward transform as follows. Let T be a $\boldsymbol{p}$-tree and let $(X_i, i \geq 1)$ be a sequence of elements of $[n]$ such that X_i is sampled according to the restriction of $\boldsymbol{p}$ to $[n] \setminus \{X_1, \dots, X_{i-1}\}$. Let $\Gamma_i = T \setminus \{X_1, \dots, X_i\}$; we stop precisely at time n, when $\{X_1, \dots, X_n\} = [n]$ and $\Gamma_n = \varnothing$.

For every $k \in [n]$, define $T_i^{\langle k \rangle}$ as the connected component of Γ_i which contains the vertex k, or $T_i^{\langle k \rangle} = \varnothing$ if $k \in \{X_1, \dots, X_i\}$. For each $i = 1, \dots, n$, let $\mathcal{U}_i$ denote the set of neighbors of X_i in Γ_{i-1}. Then we can write $\mathcal{U}_i = \{U_i^{\langle k \rangle} : 1 \le k \le n \text{ such that } T_{i-1}^{\langle k \rangle} \ni X_i\}$ where $U_i^{\langle k \rangle}$ is the unique element of $\mathcal{U}_i$ which lies in $T_i^{\langle k \rangle}$. The cuts which affect the connected component containing k are

$$\mathcal{E}_{\langle k \rangle} := \{x \in [n] : \exists i \ge 1, X_i = x \in T_{i-1}^{\langle k \rangle}\}.$$

We claim that there exists a tree G such that for every $k \in [n]$, the path $[\![X_1, k]\!]$ in G is precisely made of the nodes in $\mathcal{E}_{\langle k \rangle}$, in the order in which they appear in the permutation $(X_1, X_2, \dots, X_n)$. In the following, we write $\operatorname{cut}(T) := G$. The following proposition justifies the claim.

Proposition 4.18. *Let T be a $\boldsymbol{p}$-tree, and let V_k, $k \ge 1$, be i.i.d. $\boldsymbol{p}$-nodes, independent of T. Then, as $k \to \infty$,*

$$\operatorname{cut}(T, V_1, \dots, V_k) \xrightarrow{d} \operatorname{cut}(T).$$

Proof. We rely on the coupling we introduced in Section 4.4.2. Since, for $k \ge 1$, we have $V_1, \dots, V_k \in S_k$ and $S_k \subseteq S_{k+1}$, the tree S_k converges almost surely to a tree on $[n]$, so that $\lim_{k\to\infty} \operatorname{cut}(T; V_1, \dots, V_k)$ indeed exists with probability one. In particular, although $\operatorname{cut}(T; V_1, \dots, V_k)$ certainly depends on $V_1, \dots, V_k$, the limit only depends on the sequence $(X_i, i \ge 1)$. Indeed, $K := \inf\{k \ge 1 : [n] = \{V_1, \dots, V_k\}\}$ is a.s. finite, and for every $k \ge K$, one has $\operatorname{cut}(T; V_1, \dots, V_k) = \operatorname{cut}(T; X_1, \dots, X_n)$. We then write $\operatorname{cut}(T) := \operatorname{cut}(T; X_1, \dots, X_n)$. $\square$

Theorem 4.19 (Cut tree)**.** *Let T be a $\boldsymbol{p}$-tree on $[n]$. Then, we have $\operatorname{cut}(T) \sim \pi$.*

Proof. In the coupling defined in Section 4.4.2, we have

$$S_k = \operatorname{Span}(\operatorname{cut}(T); V_1, V_2, \dots, V_k) \to \operatorname{cut}(T)$$

almost surely as $k \to \infty$. However, by Corollary 4.15, S_k is distributed like $\operatorname{Span}(T; V_1, \dots, V_k)$, so that $S_k \to T$ in distribution, as $k \to \infty$, which completes the proof. $\square$

SHUFFLING TREES AND THE REVERSE TRANSFORMATION. Given a tree $g \in \mathbb{T}_n$ that we know is $\operatorname{cut}(t)$ for some tree $t \in \mathbb{T}_n$, and the collections of sets $\mathcal{U}_x$, $x \in [n]$, we cannot recover the initial tree t exactly, for the information about the root has been lost. However, the structure of t as an unrooted tree is easily (in this case, trivially) recovered by connecting every node x to all the nodes in $\mathcal{U}_x$. We now define the reverse operation, which samples the sets $\mathcal{U}_x$ with the correct distribution conditional on g, and produces a tree $\tilde{T}$ distributed as T conditionally on $\operatorname{cut}(T) = g$.

Consider a tree $g \in \mathbb{T}_n$, rooted at $r \in [n]$. For each edge $\langle x, w \rangle$ of the tree g, let U_w be a random element sampled according to the restriction of $\boldsymbol{p}$ to $\operatorname{Sub}(g, w)$. Let $\mathbf{U} \in \mathbb{U}(g) := \mathbb{U}(g, 1, 2, \dots, n)$ be the vector of the U_w, sorted using the canonical order on g with distinguished nodes $1, 2, \dots, n$. Let $\tau(g, [n]; \mathbf{U})$ denote the graph on $[n]$ whose edges are $\{x, U_w\}$, for $\langle x, w \rangle$ edges of g. Then, $\tau(g, [n]; \mathbf{U})$ is a tree (Lemma 4.16) and we write $\operatorname{shuff}(g)$ for the random rerooting of $\tau(g, [n]; \mathbf{U})$ at an independent $\boldsymbol{p}$-node.

Proposition 4.20. *Let G be a $\boldsymbol{p}$-tree, and $(V_k, k \ge 1)$ a sequence of i.i.d. $\boldsymbol{p}$-nodes. Then, as $k \to \infty$,*

$$\operatorname{shuff}(G; V_1, \dots, V_k) \xrightarrow{d} \operatorname{shuff}(G).$$

Proof. We prove the claim using a coupling which we build using the random variables U_w, $w \neq r$. For $k \ge 1$, we let $\mathbf{U}_k$ be the subset of $\mathbf{U}$ containing the U_w for which $w \in \operatorname{Span}^*(G; V_1, \dots, V_k)$, in the canonical order on $\operatorname{Span}^*(G; V_1, \dots, V_k)$. Then for $k \ge 1$, $\mathbf{U}_k \in \mathbb{U}(G, V_1, \dots, V_k)$ and since

$\mathrm{Span}(G; V_1, \dots, V_k)$ increases to T, the number of edges of $\tau(G; V_1, \dots, V_k; \mathbf{U}_k)$ which are constrained by the choices in $\mathbf{U}_k$ increases until they are all constrained. It follows that

$$\tau(G; V_1, \dots, V_k; \mathbf{U}_k) \to \tau(G; 1, 2, \dots, n; \mathbf{U})$$

almost surely, as $k \to \infty$. Re-rooting all the trees at the same random $\boldsymbol{p}$-node proves the claim. $\square$

We can now state the duality for the complete cutting procedure. It follows readily from the distributional identity in Proposition 4.17

$$(T, \mathrm{cut}(T, V_1, \dots, V_k)) \stackrel{d}{=} (\mathrm{shuff}(T, V_1, \dots, V_k), T).$$

and the fact that $\mathrm{cut}(T; V_1, \dots, V_k) \to \mathrm{cut}(T)$ and $\mathrm{shuff}(T; V_1, \dots, V_k) \to \mathrm{shuff}(T)$ in distribution as $k \to \infty$ (Propositions 4.18 and 4.20).

Proposition 4.21 (Cutting duality)**.** *Let T be a $\mathbf{p}$-tree. Then, we have the following duality in distribution*

$$(T, \mathrm{cut}(T)) \stackrel{d}{=} (\mathrm{shuff}(T), T).$$

In particular, $\mathrm{shuff}(T) \sim \pi$.

4.5 Cutting down an inhomogeneous continuum random tree

From now on, we fix some $\boldsymbol{\theta} = (\theta_0, \theta_1, \theta_2, \cdots) \in \boldsymbol{\Theta}$. We denote by $I = \{i \geq 1 : \theta_i > 0\}$ the index set of those θ_i with nonzero values. Let $\mathcal{T}$ be the real tree obtained from the Poisson point process construction in Section 4.2.5. We denote by μ and ℓ its respective mass and length measures. Recall the measure $\mathcal{L}$ defined by

$$\mathcal{L}(dx) = \theta_0^2 \ell(dx) + \sum_{i \in I} \theta_i \delta_{\beta_i}(dx),$$

where β_i is the branch point of local time θ_i for $i \in I$. The hypotheses on $\boldsymbol{\theta}$ entail that $\mathcal{L}$ has infinite total mass. On the other hand, we have

Lemma 4.22. *Almost surely, $\mathcal{L}$ is a σ-finite measure concentrated on the skeleton of $\mathcal{T}$. More precisely, if $(V_i, i \geq 1)$ is a sequence of independent points sampled according to μ, then for each $k \geq 1$, we have* $\mathbf{P}$*- almost surely*

$$\mathcal{L}(\mathrm{Span}(\mathcal{T}; V_1, V_2, \cdots, V_k)) < \infty.$$

Proof. We consider first the case $k = 1$. Recall the Poisson processes $(\mathrm{P}_j, j \geq 0)$ in the Section 4.2.5 and the notations there. We have seen that $\mathrm{Span}(\mathcal{T}; V_1)$ and R_1 have the same distribution. Then we have

$$\mathcal{L}(\mathrm{Span}(\mathcal{T}; V_1)) \stackrel{d}{=} \theta_0^2\, \eta_1 + \sum_{i \geq 1} \theta_i\, \delta_{\xi_{i,1}}([0, \eta_1]).$$

By construction, η_1 is either $\xi_{j,2}$ for some $j \geq 1$ or u_1. This entails that on the event $\{\eta_1 \in \mathrm{P}_j\}$, we have $\eta_1 < \xi_{i,2}$ for all $i \in \mathbb{N} \setminus \{j\}$. Then,

$$\mathbb{E}\left[\sum_{i \geq 1} \theta_i\, \delta_{\xi_{i,1}}([0, \eta_1])\right] = \sum_{j \geq 1} \mathbb{E}\left[\sum_{i \geq 1} \theta_i \cdot \mathbf{1}_{\{\xi_{i,1} \leq \eta_1\}} \mathbf{1}_{\{\eta_1 = \xi_{j,2}\}}\right] + \mathbb{E}\left[\sum_{i \geq 1} \theta_i \cdot \mathbf{1}_{\{\xi_{i,1} < \eta_1\}} \mathbf{1}_{\{\eta_1 = u_1\}}\right].$$

Note that the event $\{\xi_{j,1} \leq \eta_1\} \cap \{\eta_1 = \xi_{j,2}\}$ always occurs. By breaking the first sum on i into $\theta_j + \sum_{i\neq j} \theta_i \mathbf{1}_{\{\xi_{i,1}<\eta_1<\xi_{i,2}\}}$ and re-summing over j, we obtain

$$\begin{aligned}
\mathbb{E}\left[\sum_{i\geq 1} \theta_i \,\delta_{\xi_{i,1}}([0,\eta_1])\right] &= \sum_{j\geq 1} \theta_j \,\mathbf{P}\left(\eta_1 \in \mathrm{P}_j\right) + \sum_{j\geq 0} \mathbb{E}\left[\sum_{i\geq 1, i\neq j} \theta_i \cdot \mathbf{1}_{\{\xi_{i,1}<\eta_1<\xi_{i,2}\}} \mathbf{1}_{\{\eta_1\in \mathrm{P}_j\}}\right] \\
&= \sum_{j\geq 1} \theta_j \,\mathbf{P}\left(\eta_1 \in \mathrm{P}_j\right) + \sum_{j\geq 0}\sum_{i\neq j} \mathbb{E}[\theta_i^2 \eta_1 e^{-\theta_i \eta_1} \mathbf{1}_{\{\eta_1\in \mathrm{P}_j\}}] \\
&\leq 1 + \sum_{i\geq 1} \theta_i^2 \cdot \mathbb{E}[\eta_1],
\end{aligned}$$

where we have used the independence of $(\mathrm{P}_j, j \geq 0)$ in the second equality. The distribution of η_1 is given by (4.5). If $\theta_0 > 0$, we have $\mathbf{P}\left(\eta_1 > r\right) \leq \exp(-\theta_0^2 r^2/2)$; otherwise, we have $\mathbf{P}\left(\eta_1 > r\right) \leq (1 + \theta_1 r)e^{-\theta_1 r}$. In either case, we are able to show that $\mathbb{E}[\eta_1] < \infty$. Therefore,

$$\mathbb{E}\left[\mathcal{L}(\mathrm{Span}(\mathcal{T}; V_1))\right] = \theta_0^2 \,\mathbb{E}[\eta_1] + \mathbb{E}\left[\sum_{i\geq 1} \theta_i \,\delta_{\xi_{i,1}}([0,\eta_1])\right] < \infty.$$

In general, the variables $V_1, V_2, \cdots, V_k$ are exchangeable, therefore

$$\mathbb{E}\left[\mathcal{L}(\mathrm{Span}(\mathcal{T}; V_1, V_2, \cdots, V_k))\right] \leq k\mathbb{E}\left[\mathcal{L}(\mathrm{Span}(\mathcal{T}; V_1))\right] < \infty,$$

which proves that $\mathcal{L}$ is almost surely finite on the trees spanning finitely many random leaves. Finally, with probability one, $(V_i, i \geq 1)$ is dense in $\mathcal{T}$, thus $\mathrm{Sk}(\mathcal{T}) = \cup_{k\geq 1}]\!] r(\mathcal{T}), V_i [\![$ (see for example [10, Lemma 5]). This concludes the proof. $\square$

We recall the Poisson point process $\mathcal{P}$ of intensity measure $dt \otimes \mathcal{L}(dx)$, whose points we have used to define both the one-node-isolation procedure and the complete cutting procedure. As a direct consequence of Lemma 4.22, $\mathcal{P}$ has finitely many atoms on $[0, t] \times \mathrm{Span}(\mathcal{T}; V_1, V_2, \cdots, V_k)$ for all $t > 0$ and $k \geq 1$, almost surely. This fact will be implicitly used in the sequel.

4.5.1 An overview of the proof

Recall the hypothesis (H) on the sequence of the probability measures $(\boldsymbol{p}_n, n \geq 1)$:

$$\sigma_n = \left(\sum_{i=1}^{n} p_{ni}^2\right)^{1/2} \xrightarrow{n\to\infty} 0, \qquad \text{and} \qquad \lim_{n\to\infty} \frac{p_{ni}}{\sigma_n} = \theta_i, \qquad \text{for every } i \geq 1. \tag{H}$$

Recall the notation T^n for a $\boldsymbol{p}_n$-tree, which, from now on, we consider as a measured metric space, equipped with the graph distance and the probability measure $\boldsymbol{p}_n$. Camarri–Pitman [41] have proved that under hypothesis (H),

$$(\sigma_n T^n, \boldsymbol{p}_n) \xrightarrow[\mathrm{d,GP}]{n\to\infty} (\mathcal{T}, \mu). \tag{4.15}$$

This is equivalent to the convergence of the reduced subtrees: For each $n \geq 1$ and $k \geq 1$, write $R_k^n = \mathrm{Span}(T^n; \xi_1^n, \cdots, \xi_k^n)$ for the subtree of T^n spanning the points $\{\xi_1^n, \cdots, \xi_k^n\}$, which are k random points sampled independently with distribution $\boldsymbol{p}_n$. Similarly, let $R_k = \mathrm{Span}(\mathcal{T}; \xi_1, \ldots, \xi_k)$ be the subtree of $\mathcal{T}$ spanning the points $\{\xi_1, \cdots, \xi_k\}$, where $(\xi_i, i \geq 1)$ is an i.i.d. sequence of common law μ. Then (4.15) holds if and only if for each $k \geq 1$,

$$\sigma_n R_k^n \xrightarrow[\mathrm{d,GH}]{n\to\infty} R_k. \tag{4.16}$$

However, even if the trees converge, one expects that for the cut trees to converge, one at least needs that the measures which are used to sample the cuts also converge in a reasonable sense. Observe that $\mathcal{L}$ has an atomic part, which, as we shall see, is the scaling limit of large $\boldsymbol{p}_n$-weights. Recall that $\boldsymbol{p}_n$ is sorted: $p_{n1} \geq p_{n2} \geq \cdots p_{nn}$. For each $m \geq 1$, we denote by $\mathfrak{B}^n_m = (1, 2, \cdots, m)$ the vector of the m $\boldsymbol{p}_n$-heaviest points of T^n, which is well-defined at least for $n \geq m$. Recall that for $i \geq 1$, β_i denotes the branch point in $\mathcal{T}$ of local time θ_i, and write $\mathfrak{B}_m = (\beta_1, \beta_2, \cdots, \beta_m)$. Then Camarri and Pitman [41] also proved that

$$\left(\sigma_n T^n, \boldsymbol{p}_n, \mathfrak{B}^n_m\right) \underset{d}{\overset{n\to\infty}{\longrightarrow}} \left(\mathcal{T}, \mu, \mathfrak{B}_m\right) \tag{4.17}$$

with respect to the m-pointed Gromov–Prokhorov topology, which will allow us to prove the following convergence of the cut-measures. Let

$$\mathcal{L}_n = \sum_{i\in[n]} \frac{p_{ni}}{\sigma_n} \cdot \delta_i = \sigma_n^{-1} \boldsymbol{p}_n. \tag{4.18}$$

Recall the notation $m\restriction_A$ for the (non-rescaled) restriction of a measure to a subset A.

Proposition 4.23. *Under hypothesis* (H)*, we have*

$$\left(\sigma_n R^n_k, \mathcal{L}_n\restriction_{R^n_k}\right) \underset{d}{\overset{n\to\infty}{\longrightarrow}} \left(R_k, \mathcal{L}\restriction_{R_k}\right), \quad \forall k \geq 1, \tag{4.19}$$

with respect to the Gromov–Hausdorff–Prokhorov topology.

The proof uses the techniques developed in [15, 41] and is postponed until Section 4.7. We prove in the following subsections that the convergence in Proposition 4.23 is sufficient to entail convergence of the cut trees. To be more precise, we denote by V^n a $\boldsymbol{p}_n$-node independent of the $\boldsymbol{p}$-tree T^n, and recall that in the construction of $H^n := \mathrm{cut}(T^n, V^n)$, the node V^n ends up at the extremity of the path upon which we graft the discarded subtrees. Recall from the construction of $\mathcal{H} := \mathrm{cut}(\mathcal{T}, V)$ in Section 4.3 that there is a point U, which is at distance L_∞ from the root. In Section 4.5.2, we prove Theorem 4.4, that is: if (H) holds, then

$$(\sigma_n H^n, \boldsymbol{p}_n, V^n) \underset{d,\mathrm{GP}}{\overset{n\to\infty}{\longrightarrow}} (\mathcal{H}, \hat{\mu}, U), \tag{4.20}$$

jointly with the convergence in (4.19). From there, the proof of Theorem 4.5 is relatively short, and we provide it immediately (taking Theorem 4.4 or equivalently (4.20) for granted).

Proof of Theorem 4.5. For each $n \geq 1$, let $(\xi^n_i)_{i\geq 1}$ be a sequence of i.i.d. points of common law $\boldsymbol{p}_n$, and let $\xi^n_0 = V^n$. Let $(\xi_i)_{i\geq 1}$ be a sequence of i.i.d. points of common law $\hat{\mu}$, and let $\xi_0 = U$. We let

$$\rho_n = (\sigma_n d_{H^n}(\xi^n_i, \xi^n_j))_{i,j\geq 0} \qquad \text{and} \qquad \rho^*_n = (\sigma_n d_{H^n}(\xi^n_i, \xi^n_j))_{i,j\geq 1}$$

the distance matrices in $\sigma_n H^n = \sigma_n \, \mathrm{cut}(T^n, V_n)$ induced by the sequences $(\xi^n_i)_{i\geq 0}$ and $(\xi^n_i)_{i\geq 1}$, respectively. According to Lemma 4.10, the distribution of $\xi^n_0 = V^n$ is $\boldsymbol{p}_n$, therefore ρ_n is distributed as ρ^*_n. Writing similarly

$$\rho = (d_{\mathcal{H}}(\xi_i, \xi_j))_{i,j\geq 0} \qquad \text{and} \qquad \rho^* = (d_{\mathcal{H}}(\xi_i, \xi_j))_{i,j\geq 1},$$

where $d_{\mathcal{H}}$ denotes the distance of $\mathcal{H} = \mathrm{cut}(\mathcal{T}, V)$, (4.20) entails that $\rho_n \to \rho$ in the sense of finite-dimensional distributions. Combined with the previous argument, we deduce that ρ and ρ^* have the same distribution. However, ρ^* is the distance matrix of an i.i.d. sequence of law $\hat{\mu}$ on H^n. And the distribution of ρ determines that of V. As a consequence, the law of U is $\hat{\mu}$.

For the distribution of $(\mathcal{H}, \hat{\mu})$, it suffices to apply the second part of Lemma 4.10, which says that $(H^n, \boldsymbol{p}_n)$ is distributed like $(T^n, \boldsymbol{p}_n)$. Then comparing (4.20) with (4.15) shows that the unconditional distribution of $(\mathcal{H}, \hat{\mu})$ is that of $(\mathcal{T}, \mu)$. □

In order to prove the similar statement for the sequence of complete cut trees $G^n = \mathrm{cut}(T^n)$ that is Theorem 4.7, the construction of the limit metric space $\mathcal{G} = \mathrm{cut}(\mathcal{T})$ first needs to be justified by resorting to Aldous' theory of continuum random trees [10]. The first step consists in proving that the backbones of $\mathrm{cut}(T^n)$ converge. For each $n \geq 1$, let $(V_i^n, i \geq 1)$ be a sequence of i.i.d. points of law $\boldsymbol{p}_n$. Recall that we defined $\mathrm{cut}(\mathcal{T})$ using an increasing family $(S_k)_{k\geq 1}$, defined in (4.10). We show in Section 4.5.3 that

Lemma 4.24. *Suppose that (H) holds. Then, for each $k \geq 1$, we have*

$$\sigma_n \,\mathrm{Span}(\mathrm{cut}(T^n); V_1^n, \cdots, V_k^n) \xrightarrow[d,\mathrm{GH}]{n\to\infty} S_k, \tag{4.21}$$

jointly with the convergence in (4.19).

Combining this with the identities for the discrete trees in Section 4.4, we can now prove Theorems 4.7 and 4.8.

Proof of Theorem 4.7. By Theorem 4.19, $(\mathrm{cut}(T^n), \boldsymbol{p}_n)$ and $(T^n, \boldsymbol{p}_n)$ have the same distribution for each $n \geq 1$. Recall the notation R_k^n for the subtree of T^n spanning k i.i.d. $\boldsymbol{p}_n$-points. Then for each $k \geq 1$ we have

$$S_k^n := \mathrm{Span}(\mathrm{cut}(T^n), V_1^n, \ldots, V_k^n) \stackrel{d}{=} R_k^n.$$

Now comparing (4.21) with (4.16), we deduce immediately that, for each $k \geq 1$,

$$S_k \stackrel{d}{=} R_k.$$

In particular the family $(S_k)_{k\geq 1}$ is consistent and leaf-tight in the sense of Aldous [10]. This even holds true almost surely conditional on $\mathcal{T}$. According to Theorem 3 and Lemma 9 of [10], this entails that conditionally on $\mathrm{cut}(\mathcal{T})$, the empirical measure $\frac{1}{k}\sum_{i=1}^{k} \delta_{U_i}$ converges weakly to some probability measure ν on $\mathrm{cut}(\mathcal{T})$ such that $(U_i, i \geq 1)$ has the distribution of a sequence of i.i.d. ν-points. This proves the existence of ν. Moreover,

$$S_k \stackrel{d}{=} \mathrm{Span}(\mathrm{cut}(\mathcal{T}), \xi_1, \cdots, \xi_k),$$

where $(\xi_i, i \geq 1)$ is an i.i.d. μ-sequence. Therefore, (4.21) entails that $(\sigma_n \,\mathrm{cut}(T^n), \boldsymbol{p}_n) \to (\mathrm{cut}(\mathcal{T}), \nu)$ in distribution with respect to the Gromov–Prokhorov topology. □

Proof of Theorem 4.8. According to Theorem 3 of [10] the distribution of $(\mathrm{cut}(\mathcal{T}), \nu)$ is characterized by the family $(S_k)_{k\geq 1}$. Since S_k and R_k have the same distribution for $k \geq 1$, it follows that $(\mathrm{cut}(\mathcal{T}), \nu)$ is distributed like $(\mathcal{T}, \mu)$. □

4.5.2 Convergence of the cut-trees $\mathrm{cut}(T^n, V^n)$: Proof of Theorem 4.4

In this part we prove Theorem 4.4 taking Proposition 4.23 for granted. Let us first reformulate (4.20) in the terms of the distance matrices, which is what we actually show in the following. For each $n \in \mathbb{N}$, let $(\xi_i^n, i \geq 2)$ be a sequence of random points of T^n sampled independently according to the mass measure $\boldsymbol{p}_n$.

We set $\xi_1^n = V^n$ and let ξ_0^n be the root of $H^n = \mathrm{cut}(T^n, V^n)$. Similarly, let $(\xi_i, i \geq 2)$ be a sequence of i.i.d. μ-points and let $\xi_1 = V$. Recall that the mass measure $\hat{\mu}$ of $\mathcal{H} = \mathrm{cut}(\mathcal{T}, V)$ is defined to be the push-forward of μ by the canonical injection ϕ. We set $\widehat{\xi}_i = \phi(\xi_i)$ for $i \geq 2$, $\widehat{\xi}_1 = U$ and $\widehat{\xi}_0$ to be the root of $\mathcal{H}$.

Then the convergence in (4.20) is equivalent to the following:

$$\left(\sigma_n d_{H^n}(\xi_i^n, \xi_j^n), 0 \leq i < j < \infty\right) \xrightarrow[d]{n\to\infty} \left(d_{\mathcal{H}}(\widehat{\xi}_i, \widehat{\xi}_j), 0 \leq i < j < \infty\right), \tag{4.22}$$

jointly with

$$\left(\sigma_n d_{T^n}(\xi_i^n,\xi_j^n), 1\le i<j<\infty\right) \xrightarrow[d]{n\to\infty} \left(d_{\mathcal{T}}(\xi_i,\xi_j), 1\le i<j<\infty\right), \tag{4.23}$$

in the sense of finite-dimensional distributions. Notice that (4.23) is a direct consequence of (4.15). In order to express the terms in (4.22) with functionals of the cutting process, we introduce the following notations. For $n\in\mathbb{N}$, let $\mathcal{P}_n$ be a Poisson point process on $\mathbb{R}_+\times T^n$ with intensity measure $dt\otimes\mathcal{L}_n$, where $\mathcal{L}_n=\boldsymbol{p}_n/\sigma_n$. For $u,v\in T^n$, recall that $[\![u,v]\!]$ denotes the path between u and v. For $t\ge 0$, we denote by T_t^n the set of nodes still connected to V^n at time t:

$$T_t^n := \{x\in T^n : [0,t]\times[\![V^n,x]\!]\cap\mathcal{P}_n=\varnothing\}.$$

Recall that the remaining part of $\mathcal{T}$ at time t is $\mathcal{T}_t=\{x\in\mathcal{T}:[0,t]\times[\![V,x]\!]\cap\mathcal{P}=\varnothing\}$. We then define

$$L_t^n := \mathrm{Card}\left\{s\le t : \boldsymbol{p}_n(T_s^n)<\boldsymbol{p}_n(T_{s-}^n)\right\} \overset{a.s.}{=} \mathrm{Card}\left\{(s,x)\in\mathcal{P}_n : s\le t, x\in T_{s-}^n\right\}. \tag{4.24}$$

This is the number of cuts that affect the connected component containing V^n before time t. In particular, $L_\infty^n:=\lim_{t\to\infty}L_t^n$ has the same distribution as $L(T^n)$ in the notation of Section 4.4. Indeed, this follows from the coupling on page 107 and the fact that if $\mathcal{P}_n=\{(t_i,x_i):i\ge1\}$ such that $t_1\le t_2\le\cdots$ then (x_i) is an i.i.d. $\boldsymbol{p}_n$-sequence. Let us recall that L_t, the continuous analogue of L_t^n, is defined by $L_t=\int_0^t\mu(\mathcal{T}_s)ds$ in Section 4.3. For $n\in\mathbb{N}$ and $x\in T^n$, we define the pair $(\tau_n(x),\varsigma_n(x))$ to be the element of $\mathcal{P}_n$ separating x from V^n

$$\tau_n(x) := \inf\{t>0 : [0,t]\times[\![V^n,x]\!]\cap\mathcal{P}_n\ne\varnothing\},$$

with the convention that $\inf\varnothing=\infty$. In words, $\varsigma_n(x)$ is the first cut that appeared on $[\![V^n,x]\!]$. For $x\in\mathcal{T}$, $(\tau(x),\varsigma(x))$ is defined similarly. We notice that almost surely $\tau(\xi_j)<\infty$ for each $j\ge2$, since $\tau(\xi_j)$ is exponential with rate $\mathcal{L}([\![V,\xi_j]\!])$, which is positive almost surely. Furthermore, it follows from our construction of $H^n=\mathrm{cut}(T^n,V^n)$ that for $n\in\mathbb{N}$ and $i,j\ge2$,

$$\begin{aligned}
d_{H^n}(\xi_0^n,\xi_1^n) &= L_\infty^n-1,\\
d_{H^n}(\xi_0^n,\xi_j^n) &= L_{\tau_n(\xi_j^n)}^n-1+d_{T^n}\left(\xi_j^n,\varsigma_n(\xi_j^n)\right);\\
d_{H^n}(\xi_1^n,\xi_j^n) &= L_\infty^n-L_{\tau_n(\xi_j^n)}^n+d_{T^n}\left(\xi_j^n,\varsigma_n(\xi_j^n)\right),
\end{aligned}$$

while for $i,j\ge2$,

$$\begin{aligned}
d_{\mathcal{H}}(\widehat{\xi}_0,\widehat{\xi}_1) &= L_\infty,\\
d_{\mathcal{H}}(\widehat{\xi}_0,\widehat{\xi}_j) &= L_{\tau(\xi_j)}+d_{\mathcal{T}}\left(\xi_j,\varsigma(\xi_j)\right);\\
d_{\mathcal{H}}(\widehat{\xi}_1,\widehat{\xi}_j) &= L_\infty-L_{\tau(\xi_j)}+d_{\mathcal{T}}\left(\xi_j,\varsigma(\xi_j)\right).
\end{aligned}$$

For $n\in\mathbb{N}$ and $i,j\ge2$, if we define the event

$$\mathcal{A}_n(i,j) := \{\tau_n(\xi_i^n)=\tau_n(\xi_j^n)\} \overset{a.s.}{=} \{\varsigma_n(\xi_i^n)=\varsigma_n(\xi_j^n)\}, \tag{4.25}$$

and $\mathcal{A}_n^c(i,j)$ its complement, then on the event $\mathcal{A}_n(i,j)$, we have $d_{H^n}(\xi_i^n,\xi_j^n)=d_{T^n}(\xi_i^n,\xi_j^n)$. Similarly we define $\mathcal{A}(i,j):=\{\tau(\xi_i)=\tau(\xi_j)\}$, and note that $\mathcal{A}(i,j)=\{\varsigma(\xi_i)=\varsigma(\xi_j)\}$ almost surely. Recall that (4.15) implies that $\sigma_n d_{T^n}(\xi_i^n,\xi_j^n)\to d_{\mathcal{T}}(\xi_i,\xi_j)$. Now, on the event $\mathcal{A}_n^c(i,j)$, we have

$$d_{H^n}(\xi_i^n,\xi_j^n)=\left|L_{\tau_n(\xi_j^n)}^n-L_{\tau_n(\xi_i^n)}^n\right|+d_{T^n}\left(\xi_j^n,\varsigma_n(\xi_j^n)\right)+d_{T^n}\left(\xi_i^n,\varsigma_n(\xi_i^n)\right),$$

if $n\in\mathbb{N}$, and

$$d_{\mathcal{H}}(\widehat{\xi}_i,\widehat{\xi}_j)=\left|L_{\tau(\xi_j)}-L_{\tau(\xi_i)}\right|+d_{\mathcal{T}}\left(\xi_j,\varsigma(\xi_j)\right)+d_{\mathcal{T}}\left(\xi_i,\varsigma(\xi_i)\right),$$

for the limit case. Therefore in order to prove (4.22), it suffices to show the joint convergence of the vector

$$\Big(\mathbf{1}_{\mathcal{A}_n(i,j)}, \tau_n(\xi_i^n), \sigma_n d_{T^n}\big(\xi_j^n, \varsigma_n(\xi_j^n)\big), \big(\sigma_n L_t^n, t \in \mathbb{R}_+ \cup \{\infty\}\big)\Big)$$

to the corresponding quantities for $\mathcal{T}$, for each $i, j \geq 2$. We begin with a lemma.

Lemma 4.25. *Under* (H)*, we have the following joint convergences as* $n \to \infty$*:*

$$(\boldsymbol{p}_n(T_t^n))_{t\geq 0} \xrightarrow{d} (\mu(\mathcal{T}_t))_{t\geq 0}, \tag{4.26}$$

in Skorokhod J_1*-topology, along with*

$$\big(\mathbf{1}_{\mathcal{A}_n(i,j)}, 2 \leq i,j \leq k\big) \xrightarrow{d} \big(\mathbf{1}_{\mathcal{A}(i,j)}, 2 \leq i,j \leq k\big), \tag{4.27}$$

$$\big(\tau_n(\xi_j^n), 2 \leq j \leq k\big) \xrightarrow{d} \big(\tau(\xi_j), 2 \leq j \leq k\big), \qquad \text{and} \tag{4.28}$$

$$\big(\sigma_n d_{T^n}\big(\xi_j^n, \varsigma(\xi_j^n)\big), 2 \leq j \leq k\big) \xrightarrow{d} \big(d_{\mathcal{T}}\big(\xi_j, \varsigma_n(\xi_j)\big), 2 \leq j \leq k\big), \tag{4.29}$$

for each $k \geq 2$*, and jointly with the convergence in* (4.19).

Proof. Recall Proposition 4.23, which says that, for each $k \geq 2$

$$(\sigma_n R_k^n, \mathcal{L}_n\restriction_{R_k^n}) \xrightarrow[d]{n\to\infty} (R_k, \mathcal{L}\restriction_{R_k}),$$

in Gromov–Hausdorff–Prokhorov topology. By the properties of the Poisson point process, this entails that for $t \geq 0$,

$$(\sigma_n R_k^n, \mathcal{P}_n\restriction_{[0,t]\times R_k^n}) \xrightarrow{d} (R_k, \mathcal{P}\restriction_{[0,t]\times R_k}), \tag{4.30}$$

in Gromov–Hausdorff–Prokhorov topology, jointly with the convergence in (4.19). For each $n \in \mathbb{N}$, the pair $(\tau_n(\xi_i^n), \varsigma_n(\xi_i^n))$ corresponds to the first jump of the point process $\mathcal{P}_n$ restricted to $[\![V_1^n, \xi_i^n]\!]$. We notice that for each pair (i,j) such that $2 \leq i,j \leq k$, the event $\mathcal{A}_n(i,j)$ occurs if and only if $\tau_n(\xi_i^n \wedge \xi_j^n) \leq \min\{\tau_n(\xi_i^n), \tau_n(\xi_j^n)\}$. Similarly, $(\tau(\xi_i), \varsigma(\xi_i))$ is the first point of $\mathcal{P}$ on $\mathbb{R} \times [\![V_1, \xi_1]\!]$, and $\mathcal{A}(i,j)$ occurs if and only if $\tau(\xi_i \wedge \xi_j) \leq \min\{\tau(\xi_i), \tau(\xi_j)\}$. Therefore, the joint convergences in (4.27), (4.28) and (4.29) follow from (4.30). On the other hand, we have

$$\mathbf{1}_{\{\xi_i^n \in T_t^n\}} = \mathbf{1}_{\{t<\tau_n(\xi_i^n)\}}, \quad t \geq 0, n \geq 1$$

For each fixed $t \geq 0$, this sequence of random variables converge to $\mathbf{1}_{\{t<\tau(\xi_i)\}} = \mathbf{1}_{\{\xi_i \in \mathcal{T}_t\}}$ by (4.30). By the law of large numbers, $k^{-1}\sum_{1\leq i\leq k} \mathbf{1}_{\{t<\tau_n(\xi_j^n)\}} \to \boldsymbol{p}_n(T_t^n)$ almost surely. Then we can find a sequence $k_n \to \infty$ slowly enough such that (see also [11, Section 2.3])

$$\frac{1}{k_n}\sum_{i=1}^{k} \mathbf{1}_{\{t<\tau_n(\xi_j^n)\}} \xrightarrow{d} \mu(\mathcal{T}_t).$$

This entails that, as $n \to \infty$,

$$\boldsymbol{p}_n(T_t^n) \xrightarrow{d} \mu(\mathcal{T}_t). \tag{4.31}$$

Using (4.31) for a sequence of times $(t_m, m \geq 1)$ dense in $\mathbb{R}_+$ and combining with the fact that $t \mapsto \mu(\mathcal{T}_t)$ is decreasing, we obtain the convergence in (4.26), jointly with (4.27), (4.28), (4.29) and (4.19). □

Proposition 4.26. *Under* (H), *we have*

$$(\sigma_n L_t^n, t \geq 0) \xrightarrow[d]{n\to\infty} (L_t, t \geq 0) \tag{4.32}$$

with respect to the uniform topology, and jointly with the convergences in (4.27), (4.28) *and* (4.29). *In particular, this entails that* $L_\infty < \infty$ *almost surely. Moreover we have*

$$L_\infty \stackrel{d}{=} d_{\mathcal{T}}(r(\mathcal{T}), V), \tag{4.33}$$

where V *is a random point of distribution* μ. *The distribution of* $d_{\mathcal{T}}(r(\mathcal{T}), V)$ *is given in* (4.5).

The above proposition is a consequence of the following lemmas.

Lemma 4.27. *Jointly with* (4.27), (4.28) *and* (4.29), *we have for any* $m \geq 1$ *and* $(t_i, 1 \leq i \leq m) \in \mathbb{R}_+^m$,

$$\left(\int_0^{t_i} \boldsymbol{p}_n(T_s^n) ds,\ 1 \leq i \leq m \right) \xrightarrow[d]{n\to\infty} \left(\int_0^{t_i} \mu(\mathcal{T}_s) ds,\ 1 \leq i \leq m \right).$$

Proof. This is a direct consequence of Lemma 4.25. □

Lemma 4.28. *If we let*

$$M_t^n := \sigma_n L_t^n - \int_0^t \boldsymbol{p}_n(T_s^n) ds, \quad n \geq 1;$$

then under the hypothesis that $\sigma_n \to 0$ *as* $n \to \infty$, *the sequence of variables* $(\sigma_n M_t^n, n \geq 1)$ *converges to* 0 *in* L^2 *as* $n \to \infty$. *Moreover, this convergence is uniform on compacts.*

In particular, Lemma 4.27 and Lemma 4.28 combined entail that for any fixed $t \geq 0$, $\sigma_n L_t^n \to L_t$ in distribution. However, to obtain the convergence of $\sigma_n L_\infty^n$ to L_∞ in distribution we need the following tightness condition.

Lemma 4.29. *Under* (H), *for every* $\delta > 0$,

$$\lim_{t\to\infty} \limsup_{n\to\infty} \mathbf{P}\left(\sigma_n \left(L_\infty^n - L_t^n\right) \geq \delta\right) = 0, \tag{4.34}$$

Proof of Lemma 4.28. Let $\mathrm{N}_t^n = \mathrm{Card}\{(s, x) \in \mathcal{P}_n : s \leq t\}$ be the counting process of $\mathcal{P}_n$. Then $(\mathrm{N}_t^n, t \geq 0)$ is a Poisson process of rate $1/\sigma_n$. We write $d\mathrm{N}^n$ for the Stieltjes measure associated with $t \mapsto \mathrm{N}_t^n$. For $t \geq 0$, let

$$\mathscr{M}_t^n := L_t^n - \int_{[0,t]} \boldsymbol{p}_n(T_{s-}^n) d\mathrm{N}_s^n, \quad \text{and} \quad \mathscr{N}_t^n := \sigma_n \int_{[0,t]} \boldsymbol{p}_n(T_{s-}^n) d\mathrm{N}_s^n - \int_0^t \boldsymbol{p}_n(T_s^n) ds.$$

We notice that, by the definition of L_t^n,

$$\mathscr{M}_t^n = \sum_{(s,x) \in \mathcal{P}_n : s \leq t} \left(\mathbf{1}_{\{x \in T_{s-}^n\}} - \boldsymbol{p}_n(T_{s-}^n) \right).$$

Since $\sigma_n^{-1} \boldsymbol{p}_n = \mathcal{L}_n$, conditionally on T_{s-}^n, $\mathbf{1}_{\{x \in T_{s-}^n\}}$ is a Bernoulli random variable of mean $\boldsymbol{p}_n(T_{s-}^n)$. Therefore, we have

$$\mathbb{E}[\mathscr{M}_t^n \mid (\mathrm{N}_s^n)_{s \leq t}] = 0. \tag{4.35}$$

From this, we can readily show that $\mathscr{M}^n$ is a martingale. On the other hand, classical results on the Poisson process entail that $\mathscr{N}^n$ is also a martingale. Once combined, we see that $\mathrm{M}^n = \sigma_n \mathscr{M}^n + \mathscr{N}^n$

itself is a martingale. Therefore, by Doob's maximal inequality for the L^2-norms of martingales, we obtain for any $t \geq 0$,

$$\mathbb{E}\Big[\sup_{s\leq t}(\mathrm{M}_s^n)^2\Big] \leq 4\mathbb{E}\big[(\mathrm{M}_t^n)^2\big] = 4\mathbb{E}\big[(\sigma_n \mathscr{M}_t^n)^2\big] + 4\mathbb{E}\big[(\mathscr{N}_t^n)^2\big],$$

as a result of (4.35). Direct computation shows that

$$\mathbb{E}\big[(\mathscr{M}_t^n)^2\big] = \mathbb{E}\Big[\frac{1}{\sigma_n}\int_0^t \Big(\boldsymbol{p}_n(T_s^n) - \boldsymbol{p}_n^2(T_s^n)\Big)ds\Big], \quad \text{and} \quad \mathbb{E}\big[(\mathscr{N}_t^n)^2\big] = \mathbb{E}\Big[\sigma_n \int_0^t \boldsymbol{p}_n^2(T_s^n)ds\Big].$$

As a consequence, for any fixed t,

$$\mathbb{E}\Big[\sup_{s\leq t}(\mathrm{M}_s^n)^2\Big] \leq 4\sigma_n \mathbb{E}\Big[\int_0^t \boldsymbol{p}_n(T_s^n)ds\Big] \leq 4\sigma_n t \to 0,$$

as $n \to \infty$. □

We need an additional argument to prove Lemma 4.29. For each $n \in \mathbb{N}$ and $s \geq 0$, let $\zeta^n(s) := \inf\{t > 0 : L_t^n \geq \lfloor s \rfloor\}$ be the right-continuous inverse of L_t^n. Recall that from the construction of $H^n = \mathrm{cut}(T^n, V^n)$, there is a correspondence between the vertex sets of the remaining tree at step $\ell - 1$ and the subtree in H at X_ℓ. Then it follows Lemma 4.10 that

$$\big(\mathfrak{v}(T^n_{\zeta^n(s)}), 0 \leq s < L^n_\infty\big) \stackrel{d}{=} \big(\mathfrak{v}(\mathrm{Sub}(T^n, x_s^n)), 0 \leq s < 1 + d_{T^n}(r(T^n), V^n)\big),$$

where x_s^n is the point on the path $[\![r(T^n), V^n]\!]$ at distance $\lfloor s \rfloor$ from $r(T^n)$. In particular, this entails

$$\big(\boldsymbol{p}_n\big(T^n_{\zeta^n(s)}\big), 0 \leq s < L^n_\infty\big) \stackrel{d}{=} \big(\boldsymbol{p}_n\big(\mathrm{Sub}(T^n, x_s^n)\big), 0 \leq s < 1 + d_{T^n}(r(T^n), V^n)\big). \tag{4.36}$$

The limit of the right-hand side is easily identified using the convergence of $\boldsymbol{p}$-trees in (4.15). Combined with (4.36), this will allow us to prove Lemma 4.29 by a time-change argument.

Let V be a random point of $\mathcal{T}$ of distribution μ. For $0 \leq s \leq d_{\mathcal{T}}(r(\mathcal{T}), V)$, let x_s be the point in $[\![r(\mathcal{T}), V]\!]$ at distance s from $r(\mathcal{T})$, or $x_s = V$ if $\ell > d(r(\mathcal{T}), V)$. Similarly, we set $x_s^n = V^n$ if $s \geq 1 + d_{T^n}(r(T^n), V^n)$.

Lemma 4.30. *Under* (H)*, we have*

$$\Big(\sigma_n L^n_\infty, \big(\boldsymbol{p}_n\big(T^n_{\zeta^n(s/\sigma_n)}\big)\big)_{s\geq 0}\Big) \xrightarrow[d]{n\to\infty} \Big(d_{\mathcal{T}}(r(\mathcal{T}), V), \big(\mu(\mathrm{Sub}(\mathcal{T}, x_s))\big)_{s\geq 0}\Big),$$

where the convergence of the second coordinates is with respect to the Skorokhod J_1-topology.

Proof. Because of (4.36) and the fact $\sigma_n \to 0$, it suffices to prove that

$$\Big(\boldsymbol{p}_n\big(\mathrm{Sub}(T^n, x^n_{s/\sigma_n}\big), s \geq 0\Big) \xrightarrow[d]{n\to\infty} \Big(\mu\big(\mathrm{Sub}(\mathcal{T}, x_s\big), s \geq 0\Big),$$

with respect to the Skorokhod J_1-topology, jointly with $\sigma_n d_{T^n}(r(T^n), V^n) \to d_{\mathcal{T}}(r(\mathcal{T}), V)$ in distribution. Recall that $(\xi_i^n, i \geq 2)$ is a sequence of i.i.d. points of common law $\boldsymbol{p}_n$ and set $\xi_0^n = V^n$, $\xi_1^n = r(T^n)$, for $n \in \mathbb{N}$. Note that $(\xi_i^n, i \geq 0)$ is still an i.i.d. sequence. Then it follows from (4.15) that

$$(\sigma_n d_{T^n}(\xi_i^n, \xi_j^n), i, j \geq 0) \xrightarrow{d} (d_{\mathcal{T}}(\xi_i, \xi_j), i, j \geq 0)$$

in the sense of finite-dimensional distributions. Taking $i = 0$ and $j = 1$, we get the convergence

$$\sigma_n d_{T^n}(V^n, r(T^n)) \xrightarrow{d} d_{\mathcal{T}}(V, r(\mathcal{T})).$$

On the other hand, for $i \geq 1$, $\xi_i^n \in \mathrm{Sub}(T^n, x_s^n)$ if and only if $d_{T^n}(\xi_i^n \wedge V^n, r(T^n)) \geq s$. Since for any rooted tree (T, d, r) and $u, v \in T$ we have $2d(r, u \wedge v) = d(r, u) + d(r, v) - d(u, v)$, we deduce that for any $k, m \geq 1$ and $(s_j, 1 \leq j \leq m) \in \mathbb{R}_+^m$,

$$\left(\mathbf{1}_{\{\xi_i^n \in \mathrm{Sub}(T^n, x^n_{s_j/\sigma_n})\}}, 1 \leq i \leq k, 1 \leq j \leq m\right) \xrightarrow{d} \left(\mathbf{1}_{\{\xi_i \in \mathrm{Sub}(\mathcal{T}, x_{s_j})\}}, 1 \leq i \leq k, 1 \leq j \leq m\right),$$

jointly with $\sigma_n d_{T^n}(V^n, r(T^n)) \xrightarrow{d} d_{\mathcal{T}}(V, r(\mathcal{T}))$. Then the argument used to establish (4.31) shows the convergence of $(\boldsymbol{p}_n(\mathrm{Sub}(T^n, x^n_{s/\sigma_n}), n \geq 1)$ in the sense of finite-dimensional distributions. The convergence in the Skorokhod topology follows from the monotonicity of the function $s \mapsto \boldsymbol{p}_n(\mathrm{Sub}(T^n, x_s^n))$. $\square$

Proof of Lemma 4.29. Let us begin with a simple observation on the Skorokhod J_1-topology. Let $\mathbb{D}^\uparrow$ be the set of those functions $x : \mathbb{R}_+ \to [0, 1]$ which are nondecreasing and càdlàg. We endow $\mathbb{D}^\uparrow$ with the Skorokhod J_1-topology. Taking $\epsilon > 0$ and $x \in \mathbb{D}^\uparrow$, we denote by $\kappa_\epsilon(x) = \inf\{t > 0 : x(t) > \epsilon\}$. The following is a well-known fact. A proof can be found in [71, Ch. VI, p. 304, Lemma 2.10]

FACT If $x_n \to x$ in $\mathbb{D}^\uparrow$, $n \to \infty$ and $t \mapsto x(t)$ is strictly increasing, then $\kappa_\epsilon(x_n) \to \kappa_\epsilon(x)$ as $n \to \infty$.

If $x = (x(t), t \geq 0)$ is a process with càdlàg paths and $t_0 \in \mathbb{R}_+$, we denote by $\mathrm{R}_{t_0}[x]$ the reversed process of x at t_0:

$$\mathrm{R}_{t_0}[x](t) = x\big((t_0 - t) -\big)$$

if $t < t_0$ and $\mathrm{R}_{t_0}[x](t) = x(0)$ otherwise. For each $n \geq 1$, let $x_n(t) = \boldsymbol{p}_n(T^n_{\varsigma^n(t)})$, $t \geq 0$ and denote by $\Lambda_n = \mathrm{R}_{L^n_\infty}[x_n]$ the reversed process at L^n_∞. Similarly, let $y(t) = \mu(\mathrm{Sub}(\mathcal{T}, x_t))$, $t \geq 0$ and denote by $\Lambda = \mathrm{R}_D[y]$ for $D = d_{\mathcal{T}}(V, r(\mathcal{T}))$. Then almost surely $\Lambda_n \in \mathbb{D}^\uparrow$ for $n \in \mathbb{N}$ and $\Lambda \in \mathbb{D}^\uparrow$. Moreover, Lemma 4.30 says that

$$\big(\Lambda_n(t/\sigma_n), t \geq 0\big) \underset{d}{\overset{n\to\infty}{\longrightarrow}} \big(\Lambda(t), t \geq 0\big) \tag{4.37}$$

in $\mathbb{D}^\uparrow$. From the construction of the ICRT in Section 4.2.5 it is not difficult to show that $t \mapsto \Lambda(t)$ is strictly increasing. Then by the above FACT, we have $\sigma_n \kappa_\epsilon(\Lambda_n) \to \kappa_\epsilon(\Lambda)$ in distribution, for each $\epsilon > 0$. In particular, we have for any fixed $\delta > 0$,

$$\lim_{\epsilon \to 0} \limsup_{n \to \infty} \mathbf{P}\big(\sigma_n \kappa_\epsilon(\Lambda_n) \geq \delta\big) \leq \lim_{\epsilon \to 0} \mathbf{P}\big(\kappa_\epsilon(\Lambda) \geq \delta\big) = 0, \tag{4.38}$$

since almost surely $\Lambda(t) > 0$ for any $t > 0$.

By Lemma 4.25, the sequence $((\boldsymbol{p}_n(T^n_t))_{t \geq 0}, n \geq 1)$ is tight in the Skorokhod topology. Combined with the fact that, for each fixed n, $\boldsymbol{p}_n(T^n_t) \searrow 0$ as $t \to \infty$ almost surely, this entails that for any fixed $\epsilon > 0$

$$\lim_{t_0 \to \infty} \limsup_{n \to \infty} \mathbf{P}\left(\sup_{t \geq t_0} \boldsymbol{p}_n(T^n_t) \geq \epsilon\right) = 0. \tag{4.39}$$

Now note that if $L^n_t = k \in \mathbb{N}$, then $T^n_t = T^n_{\varsigma^n(k)}$ a.s. since no change occurs until the time of the next cut, in particular we have

$$\boldsymbol{p}_n(T^n_t) = \boldsymbol{p}_n\big(T^n_{\varsigma^n(L^n_t)}\big) \quad \text{a.s.},$$

from which we deduce that

$$\big\{\boldsymbol{p}_n(T^n_{t_0}) < \epsilon\big\} \subseteq \big\{\kappa_\epsilon(\Lambda_n) \geq L^n_\infty - L^n_{t_0}\big\} \quad \text{a.s.},$$

Then we have

$$\big\{\sigma_n\big(L^n_\infty - L^n_{t_0}\big) \geq \delta\big\} \cap \big\{\sup_{t \geq t_0} \boldsymbol{p}_n(T^n_t) < \epsilon\big\} \subseteq \big\{\sigma_n \kappa_\epsilon(\Lambda_n) \geq \delta\big\}, \quad \text{a.s.}$$

Therefore,

$$\begin{aligned}
&\limsup_{n\to\infty} \mathbf{P}\big(\sigma_n(L^n_\infty - L^n_{t_0}) \ge \delta\big)\\
&\le \limsup_{n\to\infty} \mathbf{P}\left(\sup_{t\ge t_0} \boldsymbol{p}_n(T^n_t) \ge \epsilon\right) + \limsup_{n\to\infty} \mathbf{P}\left(\sigma_n\big(L^n_\infty - L^n_{t_0}\big) \ge \delta \text{ and } \sup_{t\ge t_0} \boldsymbol{p}_n(T^n_t) < \epsilon\right)\\
&\le \limsup_{n\to\infty} \mathbf{P}\left(\sup_{t\ge t_0} \boldsymbol{p}_n(T^n_t) \ge \epsilon\right) + \limsup_{n\to\infty} \mathbf{P}\left(\sigma_n \kappa_\epsilon(\Lambda_n) \ge \delta\right).
\end{aligned}$$

In above, if we let first $t_0 \to \infty$ and then $\epsilon \to 0$, we obtain (4.34) as a combined consequence of (4.38) and (4.39). □

Proof of Proposition 4.26. We fix a sequence of $(t_m, m \ge 1)$, which is dense in $\mathbb{R}_+$. Combining Lemmas 4.27 and 4.28, we obtain, for all $k \ge 1$,

$$\big(\sigma_n L^n_{t_m}, 1 \le m \le k\big) \xrightarrow[d]{n\to\infty} \big(L_{t_m}, 1 \le m \le k\big), \tag{4.40}$$

jointly with the convergences in (4.27), (4.28), (4.29) and (4.19). We deduce from this and Lemma 4.29 that $L_\infty < \infty$ a.s. and

$$\sigma_n L^n_\infty \xrightarrow[d]{n\to\infty} L_\infty, \tag{4.41}$$

jointly with (4.27), (4.28), (4.29) and (4.19), by Theorem 4.2 of [31, Chapter 1]. Combined with the fact that $t \mapsto L_t$ is continuous and increasing, this entails the uniform convergence in (4.32). Finally, the distributional identity (4.33) is a direct consequence of Lemma 4.30. □

Proof of Theorem 4.4. We have seen that $L_\infty < \infty$ almost surely. Therefore the cut tree $(\mathrm{cut}(\mathcal{T}, V), \hat{\mu})$ is well defined almost surely. Comparing the expressions of $d_{H^n}(\xi^n_i, \xi^n_j)$ given at the beginning of this subsection with those of $d_{\mathcal{H}}(\xi_i, \xi_j)$, we obtain from Lemma 4.25 and Proposition 4.26 the convergence in (4.22). This concludes the proof. □

Remark. Before concluding this section, let us say a few more words on the proof of Proposition 4.26. The convergence of $(\sigma_n L^n_t, t \ge 0)$ to $(L_t, t \ge 0)$ on any finite interval follows mainly from the convergence in Proposition 4.23. The proof here can be easily adapted to the other models of random trees, see [30, 85]. On the other hand, our proof of the tightness condition (4.34) depends on the specific cuttings on the birthday trees, which has allowed us to deduce the distributional identity (4.36). In general, the convergence of L^n_∞ may indeed fail. An obvious example is the classical record problem (see Example 1.4 in [72]), where we have $L^n_t \to L_t$ for any fixed t, while $L^n_\infty \sim \ln n$ and therefore is not tight in $\mathbb{R}$.

4.5.3 Convergence of the cut-trees $\mathrm{cut}(T^n)$: Proof of Lemma 4.24

Let us recall the settings of the complete cutting down procedure for $\mathcal{T}$: $(V_i, i \ge 1)$ is an i.i.d. sequence of common law μ; $\mathcal{T}_{V_i}(t)$ is the equivalence class of $\sim_t$ containing V_i, whose mass is denoted by $\mu_i(t)$; and $L^i_t = \int_0^t \mu_i(s)ds$. The complete cut-tree $\mathrm{cut}(\mathcal{T})$ is defined as the complete separable metric space $\overline{\cup_k S_k}$. We introduce some corresponding notations for the discrete cuttings on T^n. For each $n \ge 1$, we sample a sequence of i.i.d. points $(V^n_i, i \ge 1)$ on T^n of distribution $\boldsymbol{p}_n$. Recall $\mathcal{P}_n$ the Poisson point process on $\mathbb{R}_+ \times T^n$ of intensity $dt \otimes \mathcal{L}_n$. We define

$$\begin{aligned}
\mu_{n,i}(t) &:= \boldsymbol{p}_n(\{u \in T^n : [0,t] \times [\![u, V^n_i]\!] \cap \mathcal{P}_n = \varnothing\}),\\
L^{n,i}_t &:= \mathrm{Card}\{s \le t : \mu_{n,i}(s) < \mu_{n,i}(s-)\}, \qquad t \ge 0, i \ge 1;\\
\tau_n(i,j) &:= \inf\{t \ge 0 : [0,t] \times [\![V^n_i, V^n_j]\!] \cap \mathcal{P}_n \ne \varnothing\}, \qquad 1 \le i,j < \infty.
\end{aligned}$$

By the construction of $G^n = \mathrm{cut}(T^n)$, we have

$$\begin{aligned} d_{G^n}(V_i^n, r(G^n)) &= L_\infty^{n,i} - 1, \\ d_{G^n}(V_i^n, V_j^n) &= L_\infty^{n,i} + L_\infty^{n,j} - 2L^{n,i}_{\tau_n(i,j)}, \quad 1 \le i, j < \infty \end{aligned} \tag{4.42}$$

where $L_\infty^{n,i} := \lim_{t\to\infty} L_t^{n,i}$ is the number of cuts necessary to isolate V_i^n. The proof of Lemma 4.24 is quite similar to that of Theorem 4.4. We outline the main steps but leave out the details.

Sketch of proof of Lemma 4.24. First, we can show with essentially the same proof of Lemma 4.25 that we have the following joint convergences: for each $k \ge 1$,

$$\big((\mu_{n,i}(t), 1 \le i \le k), t \ge 0\big) \overset{n\to\infty}{\underset{d}{\longrightarrow}} \big((\mu_i(t), 1 \le i \le k), t \ge 0\big), \tag{4.43}$$

with respect to Skorokhod J_1-topology, jointly with

$$\big(\tau_n(i,j), 1 \le i, j \le k\big) \overset{n\to\infty}{\underset{d}{\longrightarrow}} \big(\tau(i,j), 1 \le i, j \le k\big), \tag{4.44}$$

jointly with the convergence in (4.19). Then we can proceed, with the same argument as in the proof of Lemma 4.28, to showing that for any $k, m \ge 1$ and $(t_j, 1 \le j \le m) \in \mathbb{R}_+^m$,

$$\left(\int_0^{t_i} \mu_{n,i}(s)ds,\ 1 \le j \le m, 1 \le i \le k\right) \overset{n\to\infty}{\underset{d}{\longrightarrow}} \left(\int_0^{t_i} \mu_i(s)ds,\ 1 \le j \le m, 1 \le i \le k\right).$$

Since the $V_i^n, i \ge 1$ are i.i.d. $\boldsymbol{p}_n$-nodes on T^n, each process $(L_t^{n,i})_{t\ge0}$ has the same distribution as $(L_t^n)_{t\ge0}$ defined in (4.24). Then Lemmas 4.28 and 4.29 hold true for each $L^{n,i}$, $i \ge 1$. We are able to show

$$\big(\sigma_n L_t^{n,i}, 1 \le i \le k\big)_{t\ge0} \overset{n\to\infty}{\underset{d}{\longrightarrow}} \big(L_t^i, 1 \le i \le k\big)_{t\ge0}, \tag{4.45}$$

with respect to the uniform topology, jointly with the convergences (4.44) and (4.19). Comparing (4.42) with (4.10), we can easily conclude. $\square$

In general, the convergence in (4.15) does not hold in the Gromov–Hausdorff topology. However, in the case where $\mathcal{T}$ is a.s. compact and the convergence (4.15) does hold in the Gromov–Hausdorff sense, then we are able to show that one indeed has GHP convergence as claimed in Theorem 4.9. In the following proof, we only deal with the case of convergence of $\mathrm{cut}(T^n)$. The result for $\mathrm{cut}(T^n, V^n)$ can be obtained using similar arguments and we omit the details.

Proof of Theorem 4.9. We have already shown in Lemma 4.24 the joint convergence of the spanning subtrees: for each $k \ge 1$,

$$\big(\sigma_n R_k^n, \sigma_n S_k^n\big) \overset{n\to\infty}{\underset{d,\mathrm{GH}}{\longrightarrow}} \big(R_k, S_k\big). \tag{4.46}$$

We now show that for each $\epsilon > 0$,

$$\lim_{k\to\infty} \lim_{n\to\infty} \mathbf{P}\left(\max\left\{\delta_{\mathrm{GH}}(R_k^n, T^n), \delta_{\mathrm{GH}}(S_k^n, \mathrm{cut}(T^n))\right\} \ge \epsilon/\sigma_n\right) = 0. \tag{4.47}$$

Since the couples $(S_k^n, \mathrm{cut}(T^n))$ and (R_k^n, T^n) have the same distribution, it is enough to prove that for each $\epsilon > 0$,

$$\lim_{k\to\infty} \lim_{n\to\infty} \mathbf{P}\left(\sigma_n\, \delta_{\mathrm{GH}}(R_k^n, T^n) \ge \epsilon\right) = 0. \tag{4.48}$$

Let us explain why this is true when $(\sigma_n T^n, \boldsymbol{p}_n) \to (\mathcal{T}, \mu)$ in distribution in the sense of GHP. Recall the space $\mathbb{M}_c^k$ of equivalence classes of k-pointed compact metric spaces, equipped with the k-pointed Gromov–Hausdorff metric. For each $k \ge 1$ and $\epsilon > 0$, we set

$$A(k, \epsilon) := \left\{(T, d, \mathbf{x}) \in \mathbb{M}_c^k : \delta_{\mathrm{GH}}\left(T, \mathrm{Span}(T; \mathbf{x})\right) \ge \epsilon\right\}.$$

It is not difficult to check that $A(k,\epsilon)$ is a closed set of $\mathbb{M}_c^k$. Now according to the proof of Lemma 13 of [91], the mapping from $\mathbb{M}_c$ to $\mathbb{M}_c^k$: $(T,\mu)\mapsto m_k(T,A(k,\epsilon))$ is upper-semicontinuous, where $\mathbb{M}_c$ is the set of equivalence classes of compact measured metric spaces, equipped with the Gromov–Hausdorff–Prokhorov metric and m_k is defined by

$$m_k(T,A(k,\epsilon)) := \int_{T^k} \mu^{\otimes k}(d\mathbf{x})\mathbf{1}_{\{[T,\mathbf{x}]\in A(k,\epsilon)\}}.$$

Applying the Portmanteau Theorem for upper-semicontinuous mappings [31, p. 17, Problem 7], we obtain

$$\limsup_{n\to\infty} \mathbb{E}\left[m_k\big((\sigma_n T^n, \boldsymbol{p}_n), A(k,\epsilon)\big)\right] \le \mathbb{E}\left[m_k\big((\mathcal{T},\mu), A(k,\epsilon)\big)\right],$$

or, in other words,

$$\limsup_{n\to\infty} \mathbf{P}\left(\sigma_n\, \delta_{\mathrm{GH}}\left(T^n, R_k^n\right) \ge \epsilon\right) \le \mathbf{P}\left(\delta_{\mathrm{GH}}\left(\mathcal{T}, R_k\right) \ge \epsilon\right) \xrightarrow[k\to\infty]{} 0,$$

since $\delta_{\mathrm{GH}}(R_k,\mathcal{T}) \to 0$ almost surely for $\mathcal{T}$ is compact [10]. This proves (4.48) and thus (4.47). By [31, Ch. 1, Theorem 4.5], (4.46) combined with (4.47) entails the joint convergence in distribution of $(\sigma_n T^n, \sigma_n \operatorname{cut}(T^n))$ to $(\mathcal{T}, \operatorname{cut}(\mathcal{T}))$ in the Gromov–Hausdorff topology. To strengthen to the Gromov–Hausdorff–Prokhorov convergence, one can adopt the arguments in Section 4.4 of [66] and we omit the details. □

4.6 Reversing the one-cutting transformation

In this section, we justify the heuristic construction of $\operatorname{shuff}(\mathcal{H},U)$ given in Section 4.3 for an ICRT $\mathcal{H}$ and a uniform leaf U. The objective is to define formally the shuffle operation in such a way that the identity (4.12) hold. In Section 4.6.1, we rely on weak convergence arguments to justify the construction of $\operatorname{shuff}(\mathcal{H},U)$ by showing it is the limit of the discrete construction in Section 4.4.1. In Section 4.6.2, we then determine from this result the distribution of the cuts in the cut-tree $\operatorname{cut}(\mathcal{T},V)$ and prove that with the right coupling, the shuffle can yield the initial tree back (or more precisely, a tree that is in the same GHP equivalence class, which is as good as it gets).

4.6.1 Construction of the one-path reversal

Let $(\mathcal{H}, d_{\mathcal{H}}, \mu_{\mathcal{H}})$ be an ICRT rooted at $r(\mathcal{H})$, and let U be a random point in $\mathcal{H}$ of distribution $\mu_{\mathcal{H}}$. Then $\mathcal{H}$ is the disjoint union of the following subsets:

$$\mathcal{H} = \bigcup_{x\in[\![r(\mathcal{H}),U]\!]} F_x \quad \text{where} \quad F_x := \{u\in T : [\![r(\mathcal{H}),u]\!]\cap[\![r(\mathcal{H}),U]\!] = [\![r,x]\!]\}.$$

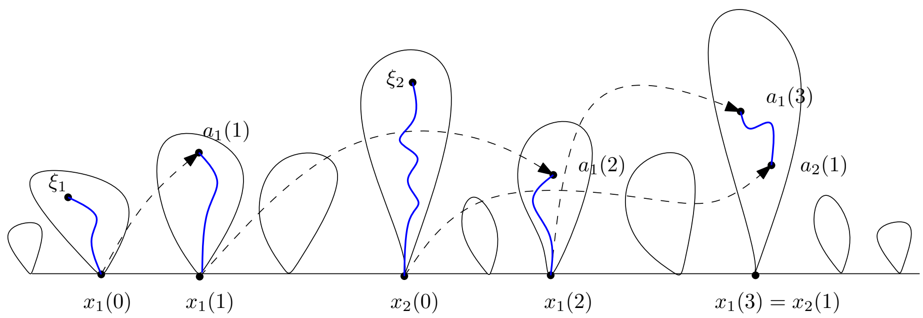

Figure 4.5 – An example with $\mathscr{I}(1,2)=3$, $\mathscr{I}(2,1)=1$ and $\mathrm{mg}(1,2)=4$. The dashed lines indicate the identifications where the root of the relevant subtrees are sent to. The blue lines represent the location of the path between ξ_1 and ξ_2 before the transformation.

It is easy to see that F_x is a subtree of T. It is nonempty ($x \in F_x$), but possibly trivial ($F_x = \{x\}$). Let $\mathbf{B} := \{x \in [\![r(\mathcal{H}), U]\!] : \mu_{\mathcal{H}}(F_x) > 0\} \cup \{U\}$, and for $x \in \mathbf{B}$, let $S_x := \mathrm{Sub}(T, x) \setminus F_x$, which is the union of those F_y such that $y \in \mathbf{B}$ and $d_{\mathcal{H}}(U, y) < d_{\mathcal{H}}(U, x)$. Then for each $x \in \mathbf{B} \setminus \{U\}$, we associate an attaching point A_x, which is independent and sampled according to $\mu_{\mathcal{H}}|_{S_x}$, the restriction of $\mu_{\mathcal{H}}$ to S_x. We also set $A_U = U$.

Now let $(\xi_i, i \geq 1)$ be a sequence of i.i.d. points of common law $\mu_{\mathcal{H}}$. The set $\boldsymbol{\mathcal{F}} := \cup_{x \in \mathbf{B}} F_x$ has full mass with probability one. Thus almost surely $\xi_i \in \boldsymbol{\mathcal{F}}$ for each $i \geq 0$. We will use $(\xi_i)_{i \geq 1}$ to span the tree $\mathrm{shuff}(\mathcal{H}, U)$ and the point ξ_1 is the future root of $\mathrm{shuff}(\mathcal{H}, U)$. For each ξ_i, we define inductively two sequences $\mathbf{x}_i := (x_i(0), x_i(1), \cdots) \in \mathbf{B}$ and $\mathbf{a}_i := (a_i(0), a_i(1), \cdots)$: we set $a_i(0) = \xi_i$, and, for $j \geq 0$,

$$x_i(j) = a_i(j) \wedge U, \qquad \text{and} \qquad a_i(j+1) = A_{x_i(j)}.$$

By definition of $(A_x, x \in \mathbf{B})$, the distance $d_{\mathcal{H}}(r(\mathcal{H}), x_i(k))$ is increasing in $k \geq 1$. For each $i, j \geq 1$, we define the merging time

$$\mathrm{mg}(i,j) := \inf\{k \geq 0 : \exists l \leq k \text{ and } x_i(l) = x_j(k-l)\},$$

with the convention $\inf \varnothing = \infty$. Another way to present $\mathrm{mg}(i,j)$ is to consider the graph on $\mathbf{B}$ with the edges $\{x, A_x \wedge U\}$, $x \in \mathbf{B}$, then $\mathrm{mg}(i,j)$ is the graph distance between $\xi_i \wedge U$ and $\xi_j \wedge U$. On the event $\{\mathrm{mg}(i,j) < \infty\}$, there is a path in this graph that has only finitely many edges, and the two walks $\mathbf{x}_i$ and $\mathbf{x}_j$ first meet at a point $y(i,j) \in \mathbf{B}$ (where by first, we mean with minimum distance to the root $r(\mathcal{H})$). In particular, if we set $\mathscr{I}(i,j), \mathscr{I}(j,i)$ to be the respective indices of the element $y(i,j)$ appearing in $\mathbf{x}_i$ and $\mathbf{x}_j$, that is,

$$\mathscr{I}(i,j) = \inf\{k \geq 0 : x_i(k) = y(i,j)\} \quad \text{and} \quad \mathscr{I}(j,i) = \inf\{k \geq 0 : x_j(k) = y(i,j)\},$$

with the convention that $\mathscr{I}(i,j) = \mathscr{I}(j,i) = \infty$ if $\mathrm{mg}(i,j) = \infty$, then $\mathrm{mg}(i,j) = \mathscr{I}(i,j) + \mathscr{I}(j,i)$. Write $\mathrm{Ht}(u) = d(u, u \wedge U)$ for the height of u in the one of F_x, $x \in \mathbf{B}$), containing it. On the event $\{\mathrm{mg}(i,j) < \infty\}$ we define $\gamma(i,j)$ which is meant to be the new distance between ξ_i and ξ_j:

$$\gamma(i,j) := \sum_{k=0}^{\mathscr{I}(i,j)-1} \mathrm{Ht}(a_i(k)) + \sum_{k=0}^{\mathscr{I}(j,i)-1} \mathrm{Ht}(a_j(k)) + d_{\mathcal{H}}(a_i(\mathscr{I}(i,j)), a_j(\mathscr{I}(j,i))),$$

with the convention if k ranges from 0 to -1, the sum equals zero.

The justification of the definition relies on weak convergence arguments: Let $\boldsymbol{p}_n$, $n \geq 1$, be a sequence of probability measures such that (H) holds with $\boldsymbol{\theta}$ the parameter of $\mathcal{H}$. Let H^n be a $\boldsymbol{p}_n$-tree and U^n a $\boldsymbol{p}_n$-node. Let $(\xi_i^n)_{i \geq 1}$ be a sequence of i.i.d. $\boldsymbol{p}_n$-points. Then, the quantities S_x^n, $\mathbf{B}^n$, $\mathbf{x}^n$, $\mathbf{a}^n$, and $\mathrm{mg}^n(i,j)$ are defined for H^n in the same way as S_x, $\mathbf{B}$, $\mathbf{x}$, $\mathbf{a}$, and $\mathrm{mg}(i,j)$ have been defined for $\mathcal{H}$. Let d_{H^n} denote the graph distance on H^n. There is only a slight difference in the definition of the distances

$$\gamma^n(i,j) := \sum_{k=0}^{\mathscr{I}^n(i,j)-1} \Big(\mathrm{Ht}(a_i^n(k)) + 1\Big) + \sum_{k=0}^{\mathscr{I}^n(j,i)-1} \Big(\mathrm{Ht}(a_j^n(k)) + 1\Big) + d_{H^n}\Big(a_i(\mathscr{I}^n(i,j)), a_j^n(\mathscr{I}^n(j,i))\Big),$$

to take into account the length of the edges $\{x, A_x^n\}$, for $x \in \mathbf{B}^n$. In that case, the sequence $\mathbf{x}^n$ (resp. $\mathbf{a}^n$) is eventually constant and equal to U^n so that $\mathrm{mg}^n(i,j) < \infty$ with probability one. Furthermore, the unique tree defined by the distance matrix $(\gamma^n(i,j) : i, j \geq 1)$ is easily seen to have the same distribution as the one defined in Section 4.4.1, since the attaching points are sampled with the same distributions and $(\gamma^n(i,j) : i, j \geq 1)$ coincides with the tree distance after attaching. Recall that we have re-rooted $\mathrm{shuff}(H^n, U^n)$ at a random point of law $\boldsymbol{p}_n$. We may suppose this point is ξ_1^n. Therefore we have (Proposition 4.12)

$$(\mathrm{shuff}(H^n, U^n), H^n) \stackrel{d}{=} (H^n, \mathrm{cut}(H^n, U^n)), \tag{4.49}$$

by Lemma 4.10.

In the case of the ICRT $\mathcal{H}$, it is a priori not clear that $\mathbf{P}(\mathrm{mg}(i,j)<\infty)=1$. We prove that

Theorem 4.31. *For any ICRT $(\mathcal{H},\mu_{\mathcal{H}})$ and a $\mu_{\mathcal{H}}$-point U, we have the following assertions:*

a) almost surely for each $i,j\geq 1$, we have $\mathrm{mg}(i,j)<\infty$;

b) almost surely the distance matrix $(\gamma(i,j),1\leq i,j<\infty)$ defines a CRT, denoted by $\mathrm{shuff}(\mathcal{H},U)$;

c) $(\mathrm{shuff}(\mathcal{H},U),\mathcal{H})$ and $(\mathcal{H},\mathrm{cut}(\mathcal{H},V))$ have the same distribution.

The main ingredient in the proof of Theorem 4.31 is the following lemma:

Lemma 4.32. *Under (H), for each $k\geq 1$, we have the following convergences*

$$\big(\sigma_n d_{H^n}(r(H^n),\mathbf{x}_i^n(j)),1\leq i\leq k,0\leq j\leq k\big)\xrightarrow[d]{n\to\infty}\big(d_{\mathcal{H}}(r(\mathcal{H}),\mathbf{x}_i(j)),1\leq i\leq k,0\leq j\leq k\big), \tag{4.50}$$

$$\big(\boldsymbol{p}_n(S^n_{x_i^n(j)}),1\leq i\leq k,0\leq j\leq k\big)\xrightarrow[d]{n\to\infty}\big(\mu_{\mathcal{H}}(S_{x_i(j)}),1\leq i\leq k,0\leq j\leq k\big), \tag{4.51}$$

and

$$\big(\sigma_n H^n,(a_i^n(j),1\leq i\leq k,0\leq j\leq k)\big)\xrightarrow[d]{n\to\infty}\big(\mathcal{H},(a_i(j),1\leq i\leq k,0\leq j\leq k)\big), \tag{4.52}$$

in the weak convergence of the pointed Gromov–Prokhorov topology.

Proof. Fix some $k\geq 1$. We argue by induction on j. For $j=0$, we note that $a_i^n(0)=\xi_i^n$ and $x_i^n(0)=\xi_i^n\wedge U^n$. Then the convergences in (4.52) and (4.50) for $j=0$ follows easily from (4.15). On the other hand, we can prove (4.51) with the same proof as in Lemma 4.30. Suppose now (4.50), (4.51) and (4.52) hold true for some $j\geq 0$. We Notice that $a_i^n(j+1)$ is independently sampled according to $\boldsymbol{p}_n$ restricted to $S_{x_i^n(j)}$, we deduce (4.52) for $j+1$ from (4.15). Then the convergence in (4.50) also follows for $j+1$, since $x_i^n(j+1)=a_i^n(j)\wedge U^n$. Finally,the very same arguments used in the proof of Lemma 4.30 show that (4.51) holds for $j+1$. $\square$

Proof of Theorem 4.31. Proof of a) By construction, $\mathrm{shuff}(H^n,U^n)$ is the reverse transformation of the one from H^n to $\mathrm{cut}(H^n,U^n)$ in the sense that each attaching "undoes" a cut. In consequence, since $\mathrm{mg}^n(i,j)$ is the number of cuts to undo in order to get ξ_i^n and ξ_j^n in the same connected component, $\mathrm{mg}^n(i,j)$ has the same distribution as the number of the cuts that fell on the path $[\![\xi_i^n,\xi_j^n]\!]$. But the latter is stochastically bounded by a Poisson variable $N_n(i,j)$ of mean $d_{H^n}(\xi_i^n,\xi_j^n)\cdot E_n(i,j)$, where $E_n(i,j)$ is an independent exponential variable of rate $d_{H^n}(U^n,\xi_i^n\wedge\xi_j^n)$. Indeed, each cut is a point of the Poisson point process $\mathcal{P}^n$ and no more cuts fall on $[\![\xi_i^n,\xi_j^n]\!]$ after the time of the first cut on $[\![U^n,\xi_i^n\wedge\xi_j^n]\!]$. But the time of the first cut on $[\![U^n,\xi_i^n\wedge\xi_j^n]\!]$ has the same distribution as $E_n(i,j)$ and is independent of $\mathcal{P}^n$ restricted on $[\![\xi_i^n,\xi_j^n]\!]$. The above argument shows that

$$\mathrm{mg}^n(i,j)=\mathscr{I}^n(i,j)+\mathscr{I}^n(j,i)\leq_{st}N_n(i,j),\quad i,j\geq 1,n\geq 1, \tag{4.53}$$

where $\leq_{st}$ denotes the stochastic domination order. It follows from (4.15) that, jointly with the convergence in (4.15), we have $N_n(i,j)\to N(i,j)$ in distribution, as $n\to\infty$, where $N(i,j)$ is a Poisson variable with parameter $d_{\mathcal{H}}(\xi_i,\xi_j)\cdot E(i,j)$ with $E(i,j)$ an independent exponential variable of rate $d_{\mathcal{H}}(U,\xi_i\wedge\xi_j)$, which is positive with probability one. Thus the sequence $(\mathrm{mg}^n(i,j),n\geq 1)$ is tight in $\mathbb{R}_+$.

On the other hand, observe that for $x \in \mathbf{B}$, $\mathbf{P}(A_x \in F_y) = \mu_{\mathcal{H}}(F_y)/\mu_{\mathcal{H}}(S_x)$ if $y \in \mathbf{B}$ and $d_{\mathcal{H}}(U, y) < d_{\mathcal{H}}(U, x)$. In particular, for two distinct points $x, x' \in \mathbf{B}$,

$$\mathbf{P}\left(\exists\, y \in \mathbf{B} \text{ such that } A_x \in F_y, A_{x'} \in F_y\right) = \sum_y \frac{\mu^2_{\mathcal{H}}(F_y)}{\mu_{\mathcal{H}}(S_x)\mu_{\mathcal{H}}(S_{x'})},$$

where the sum is over those $y \in \mathbf{B}$ such that $d_{\mathcal{H}}(U, y) < \min\{d_{\mathcal{H}}(U, x), d_{\mathcal{H}}(U, x')\}$. Similarly, for $n \geq 1$,

$$\mathbf{P}\left(\exists\, y \in \mathbf{B}^n \text{ such that } A^n_x \in F^n_y, A^n_{x'} \in F^n_y\right) = \sum_y \frac{\boldsymbol{p}^2_n(F^n_y)}{\boldsymbol{p}_n(S^n_x)\boldsymbol{p}_n(S^n_{x'})}.$$

Then it follows from (4.50) and the convergence of the masses in Lemma 4.30 that

$$\begin{aligned}\mathbf{P}\left(\mathscr{I}^n(i,j) = 1; \mathscr{I}^n(j,i) = 1\right) &= \mathbf{P}\big(\exists\, y \in \mathbf{B}^n \text{ such that } A^n_{x_i(0)} \in F^n_y, A^n_{x_j(0)} \in F^n_y\big)\\ &\overset{n\to\infty}{\longrightarrow} \mathbf{P}\left(\mathscr{I}(i,j) = 1; \mathscr{I}(j,i) = 1\right).\end{aligned}$$

By induction and Lemma 4.32, this can be extended to the following: for any natural numbers $k_1, k_2 \geq 0$, we have

$$\mathbf{P}\left(\mathscr{I}^n(i,j) = k_1; \mathscr{I}^n(j,i) = k_2\right) \overset{n\to\infty}{\longrightarrow} \mathbf{P}\left(\mathscr{I}(i,j) = k_1; \mathscr{I}(j,i) = k_2\right).$$

Combined with the tightness of $(\mathrm{mg}^n(i,j), n \geq 1) = (\mathscr{I}^n(i,j) + \mathscr{I}^n(j,i), n \geq 1)$, this entails that

$$(\mathscr{I}^n(i,j), \mathscr{I}^n(j,i)) \underset{d}{\overset{n\to\infty}{\longrightarrow}} (\mathscr{I}(i,j), \mathscr{I}(j,i)), \quad i,j \geq 1 \tag{4.54}$$

jointly with (4.50) and (4.52), using the usual subsequence arguments. In particular, $\mathscr{I}(i,j) + \mathscr{I}(j,i) \leq_{st} N(i,j) < \infty$ almost surely, which entails that $\mathrm{mg}(i,j) < \infty$ almost surely, for each pair $(i,j) \in \mathbb{N}\times\mathbb{N}$.

Proof of b) It follows from (4.52), (4.54) and the expression of $\gamma(i,j)$ that

$$\big(\sigma_n \gamma^n(i,j), i,j \geq 1\big) \underset{d}{\overset{n\to\infty}{\longrightarrow}} \big(\gamma(i,j), i,j \geq 1\big), \tag{4.55}$$

in the sense of finite-dimensional distributions, jointly with the Gromov–Prokhorov convergence of $\sigma_n H^n$ to $\mathcal{H}$ in (4.15). However by (4.49), the distribution of $\mathrm{shuff}(H^n, U^n)$ is identical to H^n. Hence, the unconditional distribution of $(\gamma(i,j), 1 \leq i,j < \infty)$ is that of the distance matrix of the ICRT $\mathcal{H}$. We can apply Aldous' CRT theory [10] to conclude that for a.e. $\mathcal{H}$, the distance matrix $(\gamma(i,j), i,j \geq 1)$ defines a CRT, denoted by $\mathrm{shuff}(\mathcal{H}, U)$. Moreover, there exists a mass measure $\tilde{\mu}$, such that if $(\tilde{\xi}_i)_{i\geq 1}$ is an i.i.d. sequence of law $\tilde{\mu}$, then

$$(d_{\mathrm{shuff}(\mathcal{H},U)}(\tilde{\xi}_i, \tilde{\xi}_j), 1 \leq i,j < \infty) \overset{d}{=} (\gamma(i,j), 1 \leq i,j < \infty).$$

Therefore, we can rewrite (4.55) as

$$\big(\sigma_n \,\mathrm{shuff}(H^n, U^n), \sigma_n H^n\big) \underset{d}{\overset{n\to\infty}{\longrightarrow}} \big(\mathrm{shuff}(\mathcal{H}, U), \mathcal{H}\big), \tag{4.56}$$

with respect to the Gromov–Prokhorov topology.

Proof of c) This is an easy consequence of (4.49) and (4.56). Let f, g be two arbitrary bounded functions continuous in the Gromov–Prokhorov topology. Then (4.56) and the continuity of f, g entail that

$$\begin{aligned}\mathbb{E}\left[f(\mathrm{shuff}(\mathcal{H}, U)) \cdot g(\mathcal{H})\right] &= \lim_{n\to\infty} \mathbb{E}\left[f(\sigma_n \,\mathrm{shuff}(H^n, U^n)) \cdot g(\sigma_n H^n)\right]\\ &= \lim_{n\to\infty} \mathbb{E}\left[f(\sigma_n H^n) \cdot g(\sigma_n \,\mathrm{cut}(H^n, U^n))\right]\\ &= \mathbb{E}\left[f(\mathcal{H}) \cdot g(\mathrm{cut}(\mathcal{H}, U))\right],\end{aligned}$$

where we have used (4.49) in the second equality. Thus we obtain the identity in distribution in c). $\square$

4.6.2 Distribution of the cuts

According to Proposition 4.12 and (4.49), the attaching points $a_i^n(j)$ have the same distribution as the points where the cuts used to be connected to in the $\boldsymbol{p}_n$ tree H^n. Then Theorem 4.31 suggests that the weak limit $a_i(j)$ should play a similar role for the continuous tree. Indeed in this section, we show that $a_i(j)$ represent the "holes" left by the cutting on $(\mathcal{T}_t)_{t\geq 0}$.

Let $(\mathcal{T}, d_{\mathcal{T}}, \mu)$ be the ICRT in Section 4.5. The μ-point V is isolated by successive cuts, which are elements of the Poisson point process $\mathcal{P}$. Now let ξ_1', ξ_2' be two independent points sampled according to μ. We plan to give a description of the image of the path $[\![\xi_1', \xi_2']\!]$ in the cut tree $\mathrm{cut}(\mathcal{T}, V)$, which turns out to be dual to the construction of one path in $\mathrm{shuff}(\mathcal{H}, U)$.

During the cutting procedure which isolates V, the path $[\![\xi_1', \xi_2']\!]$ is pruned from the two ends into segments. See Figure 4.6. Each segment is contained in a distinct portion $\Delta\mathcal{T}_t := \mathcal{T}_{t-} \setminus \mathcal{T}_t$, which is discarded at time t. Also recall that $\Delta\mathcal{T}_t$ is grafted on the interval $[0, L_\infty]$ to construct $\mathrm{cut}(\mathcal{T}, V)$. The following is just a formal reformulation:

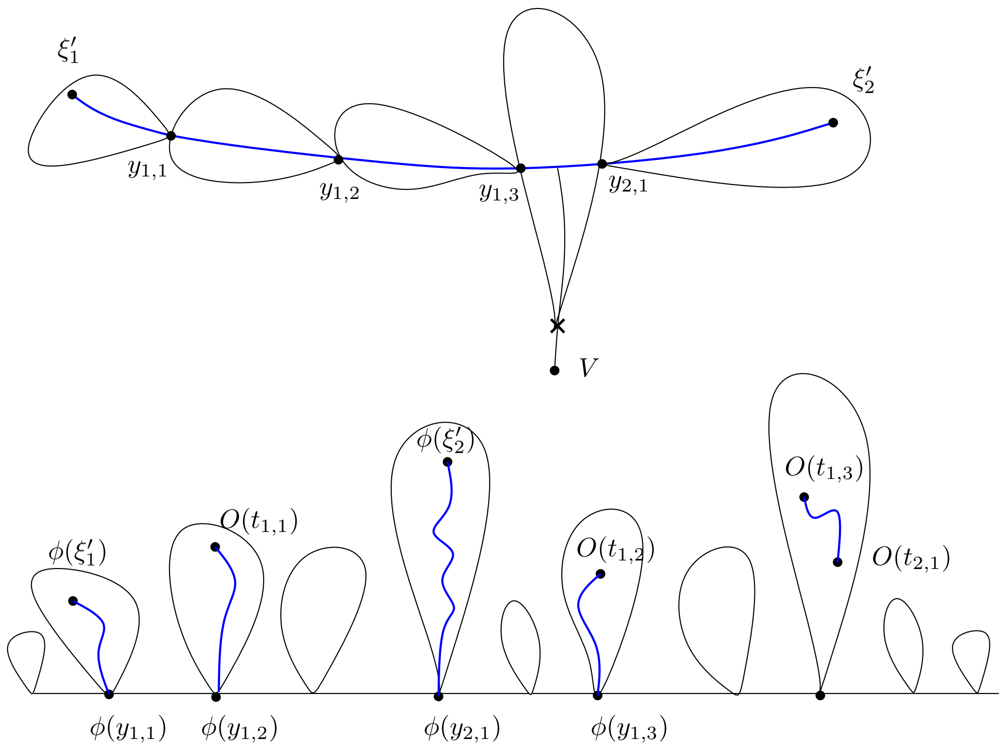

Figure 4.6 – *An example with $M_1 = 3$ and $M_2 = 1$. Above, the cuts partition the path between ξ_1' and ξ_2' into segments. The cross represents the first cut on $[\![\xi_1', V]\!] \cap [\![\xi_2', V]\!]$. Below, the image of these segments in* $\mathrm{cut}(\mathcal{T}, V)$.

Lemma 4.33. *Let*

$$(t_{1,1}, y_{1,1}), (t_{1,2}, y_{1,2}), \cdots, (t_{1,M_1}, y_{1,M_1}) \qquad \textit{and} \qquad (t_{2,1}, y_{2,1}), (t_{2,2}, y_{2,2}), \cdots, (t_{2,M_2}, y_{2,M_2}),$$

be the respective (finite) sequences of cuts on $[\![\xi_1', V]\!] \cap [\![\xi_1', \xi_2']\!]$ and $[\![\xi_2', V]\!] \cap [\![\xi_1', \xi_2']\!]$ such that $0 < t_{i,1} < t_{i,2} < \cdots < t_{i,M_i} < \infty$ for $i = 1, 2$. Then the points $\{y_{i,j} : 1 \leq j \leq M_i, i = 1, 2\}$ partition of the path $[\![\xi_1', \xi_2']\!]$ into segments and:

- *for $i = 1, 2$, $[\![\xi_i', y_{i,1}]\!] \subset \Delta\mathcal{T}_{t_{i,1}}$;*
- *for $j = 1, 2, \cdots, M_i - 2$, $]\!]y_{i,j}, y_{i,j+1}]\!] \subset \Delta\mathcal{T}_{t_{i,j+1}}$.*

Finally, writing

$$t_{me} := \inf\{t > 0 : \mathcal{P}_t \cap [\![\xi'_1, V]\!] \cap [\![\xi'_2, V]\!] \neq \varnothing\} < \infty,$$

$]\!]y_{1,M_1}, y_{2,M_2}[\![$ *is contained in* $\Delta\mathcal{T}_{t_{me}}$.

Proof. It suffices to prove that M_1, M_2 are finite with probability 1. The other statements are straightforward from the cutting procedure. But an argument similar to the one used in the proof of a) of Theorem 4.31 shows that $M_1 + M_2$ is stochastically bounded by a Poisson variable with mean $d_{\mathcal{T}}(\xi'_1, \xi'_2) \cdot t_{me}$, which entails that M_1, M_2 are finite almost surely. □

Recall that $\mathrm{cut}(\mathcal{T}, V)$ is defined so as to be a complete metric space. Denote by ϕ the canonical injection from $\cup_{t\in\mathcal{C}}\Delta\mathcal{T}_t$ to $\mathrm{cut}(\mathcal{T}, V)$. For $1 \leq j \leq M_i - 2$ and $i = 1, 2$, it is not difficult to see that there exists some point $O(t_{i,j})$ of $\mathrm{cut}(\mathcal{T}, V)$ such that the closure of $\phi(]\!]y_{i,j}, y_{i,j+1}]\!])$ is $[\![O(t_{i,j}), \phi(y_{i,j+1})]\!]$. Similarly, the closure of $\phi(]\!]y_{1,M_1}, y_{2,M_2}[\![)$ is equal to $[\![O(t_{1,M_1}), O(t_{2,M_2})]\!]$, with $O(t_{1,M_1}), O(t_{2,M_2})$ two leaves contained in the closure of $\phi(\Delta\mathcal{T}_{t_{me}})$. Comparing this with Theorem 4.31, one may suspect that $\{O(t_{1,j}) : 1 \leq j \leq M_1\}, \{O(t_{2,j}) : 1 \leq j \leq M_2\}$ should have the same distribution as $\{a_1(j) : 1 \leq j \leq \mathscr{I}(1,2)\}, \{a_2(j) : 1 \leq j \leq \mathscr{I}(2,1)\}$. This is indeed true. In the following, we show a slightly more general result about all the points. For each $t \in \mathcal{C} = \{t > 0 : \mu(\Delta\mathcal{T}_t) > 0\}$, let $x(t) \in \mathcal{T}$ be the point such that $(t, x(t)) \in \mathcal{P}$. Then we can define $O(t)$ to be the point of $\mathrm{cut}(\mathcal{T}, V)$ which marks the "hole" left by the cutting at $x(t)$. More precisely, let (t', x') be the first element after time t of $\mathcal{P}$ on $[\![r(\mathcal{T}), x(t)[\![$. Then there exists some point $O(t)$ such that the closure of $\phi(]\!]x(t), x']\!])$ in $\mathrm{cut}(\mathcal{T}, V)$ is $[\![O(t), \phi(x')]\!]$.

Proposition 4.34. *Conditionally on* $\mathrm{cut}(\mathcal{T}, V)$*, the collection* $\{O(t), t \in \mathcal{C}\}$ *is independent, and each* $O(t)$ *has distribution* $\hat{\mu}$ *restricted to* $\cup_{s>t}\phi(\mathcal{T}_{s-} \setminus \mathcal{T}_s)$.

Proof. It suffices to show that $\{O(t), t \in \mathcal{C}\}$ has the same distribution as the collection of attaching points $\{A_x, x \in \mathbf{B}\}$ introduced in the previous section. Observe that if we take $(\mathcal{H}, U) = (\mathrm{cut}(\mathcal{T}, V), L_\infty)$ and replace $\{A_x, x \in \mathbf{B}\}$ with $\{O(t), t \in \mathcal{C}\}$, then it follows that $\mathrm{shuff}(\mathcal{H}, U)$ is isometric to $\mathcal{T}$, since the two trees are metric completions of the same distance matrix with probability one. In particular, we have

$$(\mathrm{shuff}(\mathcal{H}, U), \mathcal{H}) \stackrel{d}{=} (\mathcal{T}, \mathrm{cut}(\mathcal{T}, V)). \tag{4.57}$$

Therefore, to determine the distribution of $\{O(t), t \in \mathcal{C}\}$, we only need to argue that the distribution of $\{A_x, x \in \mathbf{B}\}$ is the unique distribution for which (4.57) holds. To see this, we notice that (4.57) implies that the distribution of $(\gamma(i,j))_{i,j\geq 1}$ is unique. But from the distance matrix $(\gamma(i,j))_{i,j\geq 1}$ (also given $\mathcal{H}$ and $(\xi_i, i \geq 1)$), we can recover $(a_i(1), i \geq 1)$, which is a size-biased resampling of $(A_x, x \in \mathbf{B})$. Indeed, the sequence $(\xi_k)_{k\geq 1}$ is everywhere dense in $\mathcal{H}$. For $x \in \mathbf{B}$, let $(\xi_{m_k}, k \geq 1)$ be the subsequence consisting of the ξ_i contained in F_x. Then $a_i(1) \in F_x$ if and only if $\liminf_{k\to\infty} \gamma(i, m_k) - \mathrm{Ht}(\xi_i) = 0$, where $\mathrm{Ht}(\xi_i) = d_{\mathcal{H}}(\xi_i, \xi_i \wedge U)$. Moreover, if the latter holds, we also have $d_{\mathcal{H}}(a_i(1), \xi_{m_k}) = \gamma(i, m_k) - \mathrm{Ht}(\xi_i)$. By Gromov's reconstruction theorem [64, $3\frac{1}{2}$], we can determine $a_i(1)$ for each $i \geq 1$. By the previous arguments, this concludes the proof. □

The above proof also shows that if we use $(O(t), t \in \mathcal{C})$ to define the points $(A_x, x \in \mathbf{B})$ then the shuffle operation yields a tree that is undistinguishable from the original ICRT $\mathcal{T}$.

4.7 Convergence of the cutting measures: Proof of Proposition 4.23

Recall the setting at the beginning of Section 4.5.1. Then proving Proposition 4.23 amounts to show that for each $k \geq 1$, we have

$$\left(\sigma_n R_k^n, \mathcal{L}_n\restriction_{R_k^n}\right) \xrightarrow[d]{n\to\infty} \left(R_k(\mathcal{T}), \mathcal{L}\restriction_{R_k}\right) \tag{4.58}$$

in Gromov–Hausdorff–Prokhorov topology. Observe that the Gromov–Hausdorff convergence is clear from (4.15), so that it only remains to prove the convergence of the measures.

Case 1. We first prove the claim assuming that $\theta_i > 0$ for every $i \geq 0$. In this case, define

$$m_n := \min\left\{ j : \sum_{i=1}^{j} \left(\frac{p_{ni}}{\sigma_n}\right)^2 \geq \sum_{i\geq 1} \theta_i^2 \right\},$$

and observe that $m_n < \infty$ since $\sum_{i\leq n}(p_{ni}/\sigma_n)^2 = 1 \geq \sum_{i\geq 1}\theta_i^2$. Note also that $m_n \to \infty$. Indeed, for every integer $k \geq 1$, since $p_{ni}/\sigma_n \to \theta_i$, for $i \geq 1$, and $\theta_{k+1} > 0$, we have, for all n large enough,

$$\sum_{i=1}^{k} \left(\frac{p_{ni}}{\sigma_n}\right)^2 < \sum_{i\geq 1} \theta_i^2,$$

so that $m_n > k$ for all n large enough. Furthermore $\lim_{j\to\infty}\theta_j = 0$, and (H) implies that

$$\lim_{n\to\infty} \frac{p_{m_n}}{\sigma_n} = 0. \tag{4.59}$$

Combining this with the definition of m_n, it follows that, as $n \to \infty$,

$$\sum_{i\leq m_n} \left(\frac{p_{ni}}{\sigma_n}\right)^2 \to \sum_{i\geq 1} \theta_i^2. \tag{4.60}$$

If $n, k, M \geq 1$, we set

$$\mathcal{L}_n^* = \sum_{m_n < i \leq n} \frac{p_{ni}}{\sigma_n}\delta_i, \quad \text{ and } \quad \Sigma(n,k,M) = \sum_{M < i \leq m_n} \frac{p_{ni}}{\sigma_n}\mathbf{1}_{\{i\in R_k^n\}}.$$

Let ℓ_n denote the (discrete) length measure on T^n. Clearly, $\sigma_n\ell_n$ is the length measure of the rescaled tree $\sigma_n T^n$, seen as a real tree.

Lemma 4.35. *Suppose that* (H) *holds. Then, for each $k \geq 1$, we have the following assertions:*

a) as $n \to \infty$, in probability

$$\delta_{\mathrm{P}}\left(\mathcal{L}_n^*\restriction_{R_k^n}, \theta_0^2\sigma_n\ell_n\restriction_{R_k^n}\right) \to 0; \tag{4.61}$$

b) for each $\epsilon > 0$, there exists $M = M(k,\epsilon) \in \mathbb{N}$ such that

$$\limsup_{n\to\infty} \mathbf{P}\big(\Sigma(n,k,M) \geq \epsilon\big) \leq \epsilon; \tag{4.62}$$

Before proving Lemma 4.35, let us first explain why this entails Proposition 4.23.

Proof of Proposition 4.23 in the Case 1. By Skorokhod representation theorem and a diagonal argument, we can assume that the convergence $(\sigma_n T^n, \mu_n, \mathfrak{B}_m^n) \to (\mathcal{T}, \mu, \mathfrak{B}_m)$, holds almost surely in the m-pointed Gromov–Prokhorov topology for all $m \geq 1$. Since the length measure ℓ_n (resp. ℓ) depends continuously on the metric of T^n (resp. the metric of $\mathcal{T}$), according to Proposition 2.23 of [85] this implies that, for each $k \geq 1$,

$$\left(\sigma_n R_k^n, \theta_0^2\sigma_n\ell_n\restriction_{R_k^n}\right) \to \left(R_k, \theta_0^2\ell\restriction_{R_k}\right), \tag{4.63}$$

almost surely in the Gromov–Hausdorff–Prokhorov topology. On the other hand, we easily deduce from the convergence of the vector $\mathfrak{B}_m^n$ and (H) that, for each fixed $m \geq 1$,

$$\left(\sigma_n R_k^n, \sum_{i=1}^{m} \frac{p_{ni}}{\sigma_n}\delta_i\restriction_{R_k^n}\right) \to \left(R_k, \sum_{i=1}^{m} \theta_i\delta_{\mathcal{B}_i}\restriction_{R_k}\right), \tag{4.64}$$

almost surely in the Gromov–Hausdorff–Prokhorov topology. In the following, we write $\delta_{\mathrm{P}}^{\mathrm{n,k}}$ (resp. $\delta_{\mathrm{P}}^{\mathrm{k}}$) for the Prokhorov distance on the finite measures on the set R_k^n (resp. R_k). In particular, since the measures below are all restricted to either R_k^n or R_k, we omit the notations $\upharpoonright_{R_k^n}$, $\upharpoonright_{R_k}$ when the meaning is clear from context. We write

$$\mathrm{Kt}_m(\mathcal{L}) := \theta_0^2 \ell + \sum_{i=1}^{m} \theta_i \delta_{\mathcal{B}_i}$$

for the cut-off measure of $\mathcal{L}$ at level m. By Lemma 4.22, the restriction of $\mathcal{L}$ to R_k^n is a finite measure. Therefore, $\mathrm{Kt}_m(\mathcal{L}) \to \mathcal{L}$ almost surely in $\delta_{\mathrm{P}}^{\mathrm{k}}$ as $m \to \infty$.

Now fix some $\epsilon > 0$. By Lemma 4.35 we can choose some $M = M(k, \epsilon)$ such that (4.62) holds, as well as

$$\mathbb{P}(\delta_{\mathrm{P}}^{\mathrm{k}}(\mathrm{Kt}_M(\mathcal{L}), \mathcal{L}) \geq \epsilon) \leq \epsilon. \tag{4.65}$$

Define now the approximation

$$\vartheta_{n,M} := \theta_0^2 \sigma_n \ell_n + \sum_{i \leq M} \frac{p_{ni}}{\sigma_n} \delta_i.$$

Then recalling the definition of $\mathcal{L}_n$ in (4.18), and using (4.62) and (4.61), we obtain

$$\limsup_{n\to\infty} \mathbb{P}(\delta_{\mathrm{P}}^{\mathrm{n,k}}(\vartheta_{n,M}, \mathcal{L}_n) \geq \epsilon) \leq \epsilon. \tag{4.66}$$

We notice that

$$\big(\sigma_n R_k^n, \vartheta_{n,M}\big) \to \big(R_k, \mathrm{Kt}_M(\mathcal{L})\big) \tag{4.67}$$

almost surely in the Gromov–Hausdorff–Prokhorov topology as a combined consequence of (4.63) and (4.64). Finally, by the triangular inequality, we deduce from (4.65), (4.66) and (4.67) that

$$\limsup_{n\to\infty} \mathbf{P}\left(\delta_{\mathrm{GHP}}\left((\sigma_n R_k^n, \mathcal{L}_n), (R_k, \mathcal{L})\right) \geq 2\epsilon\right) \leq 2\epsilon,$$

for any $\epsilon > 0$, which concludes the proof. □

Proof of Lemma 4.35. We first consider the case $k = 1$. Define

$$D_n := d_{T^n}\big(r(T^n), V_1^n\big), \qquad \text{and} \qquad F_n^{\mathcal{L}}(l) := \mathcal{L}_n^*\big(\mathbf{B}(r(T^n), l) \cap R_1^n\big), \tag{4.68}$$

where $\mathbf{B}(x, l)$ denotes the ball in T^n centered at x and with radius l. Then the function $F_n^{\mathcal{L}}$ determines the measure $\mathcal{L}_n^* \upharpoonright_{R_1^n}$ in the same way a distributional function determines a finite measure of $\mathbb{R}_+$. Let $(X_j^n, j \geq 0)$ be a sequence of i.i.d. random variables of distribution $\boldsymbol{p}_n$. We define $\mathfrak{R}_0^n = 0$, and for $m \geq 1$,

$$\mathfrak{R}_m^n = \inf\left\{j > \mathfrak{R}_{m-1}^n : X_j^n \in \{X_1^n, X_2^n, \cdots, X_{j-1}^n\}\right\}$$

the m-th repeat time of the sequence. For $l \geq 0$, we set

$$F_n(l) := \sum_{j=0}^{l \wedge (\mathfrak{R}_1^n - 1)} \sum_{i > m_n} \frac{p_{ni}}{\sigma_n} \mathbf{1}_{\{X_j^n = i\}}.$$

According to the construction of the birthday tree in [41] and Corollary 3 there, we have

$$\big(D_n, F_n^{\mathcal{L}}(\cdot)\big) \stackrel{d}{=} \big(\mathfrak{R}_1^n - 1, F_n(\cdot)\big). \tag{4.69}$$

Let $q_n \geq 0$ be defined by $q_n^2 = \sum_{i > m_n} p_{ni}^2$. Then (4.60) entails $\lim_{n\to\infty} q_n/\sigma_n = \theta_0$. For $l \geq 0$, we set

$$Z_n(l) := \left| F_n(l) - \frac{q_n^2}{\sigma_n}\big((l+1) \wedge \mathfrak{R}_1^n\big) \right|.$$

We claim that $\sup_{l\geq 0} Z_n(l) \to 0$ in probability as $n\to\infty$. To see this, observe first that

$$Z_n(l) = \left| \sum_{j=0}^{l\wedge(\mathfrak{R}_1^n-1)} \left(\sum_{i>m_n} \frac{p_{ni}}{\sigma_n} \mathbf{1}_{\{X_j^n=i\}} - \frac{q_n^2}{\sigma_n} \right) \right|,$$

where the terms in the parenthesis are variables which are independent, centered, and of variance $\chi_n := \sigma_n^{-2}\sum_{i>m_n} p_{ni}^3 - \sigma_n^{-2} q_n^4$. Therefore, Doob's maximal inequality entails that for any fixed number $N>0$,

$$\begin{aligned}
\mathbb{E}\left[\left(\sup_{l\geq 0} Z_n(l) \mathbf{1}_{\{\mathfrak{R}_1^n\leq N/\sigma_n\}} \right)^2 \right] &\leq \mathbb{E}\left[\left(\sup_{l<\lfloor N/\sigma_n\rfloor} \sum_{j=0}^{l} \left(\sum_{i>m_n} \frac{p_{ni}}{\sigma_n} \mathbf{1}_{\{X_j^n=i\}} - \frac{q_n^2}{\sigma_n} \right) \right)^2 \right] \\
&\leq 4N\sigma_n^{-1}\chi_n \\
&\leq 4N \frac{q_n^2}{\sigma_n^2} \frac{p_{nm_n}+q_n^2}{\sigma_n} \to 0
\end{aligned}$$

by (4.59) and the fact that $q_n/\sigma_n \to \theta_0$. In particular, it follows that

$$\sup_{l\geq 0} Z_n(l) \mathbf{1}_{\{\mathfrak{R}_1^n\leq N/\sigma_n\}} \to 0, \tag{4.70}$$

in probability as $n\to\infty$. On the other hand, the convergence of the $\boldsymbol{p}_n$-trees in (4.15) implies that the family of distributions of $(\sigma_n D_n, n\geq 1)$ is tight. By (4.69), this entails that

$$\lim_{N\to\infty} \limsup_{n\to\infty} \mathbf{P}\left(\mathfrak{R}_1^n > N/\sigma_n\right) = 0. \tag{4.71}$$

Combining this with (4.70) proves the claim.

The generalized distribution function as in (4.68) for the discrete length measure ℓ_n is $l\mapsto l\wedge D_n$. Thus, since $\sup_l Z_n(l)\to 0$ in probability, the identity in (4.69) and $q_n/\sigma_n\to\theta_0$ imply that

$$\delta_{\mathrm{P}}\left(\mathcal{L}_n^*\restriction_{R_1^n}, \theta_0^2\sigma_n\ell_n\restriction_{R_1^n} \right) \to 0$$

in probability as $n\to\infty$. This is exactly (4.61) for $k=1$.

In the general case where $k\geq 1$, we set

$$D_{n,1} := D_n, \qquad D_{n,m} := d_{T^n}\big(b_n(m), V_m^n\big), \quad m\geq 2,$$

where $b_n(m)$ denotes the branch point of T^n between V_m^n and R_{m-1}^n, i.e, $b_n(m)\in R_{m-1}^n$ such that $[\![r(T^n), V_m^n]\!] \cap R_{m-1}^n = [\![r(T^n), b_n(m)]\!]$. We also define

$$F_{n,1}^{\mathcal{L}}(l) := F_n^{\mathcal{L}}, \qquad \text{and} \qquad F_{n,m}^{\mathcal{L}}(l) := \mathcal{L}_n^*\big(\mathbf{B}(b_n(m),l)\cap]\!]b_n(m),V_m^n]\!]\big), \quad m\geq 2.$$

Then conditional on $\mathfrak{R}_k^n$, the vector $(F_{n,1}^{\mathcal{L}}(\cdot),\cdots,F_{n,k}^{\mathcal{L}}(\cdot))$ determines the measure $\mathcal{L}_n^*\restriction_{R_k^n}$ for the same reason as before. If we set

$$F_{n,1}(l) := F_n(l), \qquad \text{and} \qquad F_{n,m}(l) := \sum_{j=\mathfrak{R}_{m-1}^n+1}^{l\wedge(\mathfrak{R}_m^n-1)} \sum_{i>m_n} \frac{p_{ni}}{\sigma_n}\mathbf{1}_{\{X_j^n=i\}}, \quad m\geq 2,$$

then Corollary 3 of [41] entails the equality in distribution

$$\big((D_{n,m}, F_{n,m}^{\mathcal{L}}(\cdot)), 1\leq m\leq k\big) \stackrel{d}{=} \big((\mathfrak{R}_m^n-\mathfrak{R}_{m-1}^n-1, F_{n,m}(\cdot)), 1\leq m\leq k\big)$$

Then by the same arguments as before we can show that

$$\max_{1\le m\le k}\sup_{l\ge 0}\left|F_{n,m}(l)-\frac{q_n^2}{\sigma_n}\Big(l\wedge(\mathfrak{R}_m^n-\mathfrak{R}_{m-1}^n-1)\Big)\right|\to 0$$

in probability as $n\to\infty$. This then implies (4.61) by the same type of argument as before.

Now let us consider (4.62). The idea is quite similar. For each $M\ge 1$, we set

$$\tilde{Z}_{n,M}:=\sum_{j=0}^{\mathfrak{R}_1^n-1}\sum_{M<i\le m_n}\frac{p_{ni}}{\sigma_n}\mathbf{1}_{\{X_j^n=i\}}.$$

Then

$$\mathbb{E}\left[\tilde{Z}_{n,M}\mathbf{1}_{\{\mathfrak{R}_1^n\le N/\sigma_n\}}\right]\le N\bigg(\sum_{M<i\le m_n}\frac{p_{ni}^2}{\sigma_n^2}\bigg).$$

Using (4.59), (H) and the fact that $\sum_i\theta_i^2<\infty$, we can easily check that for any fixed N,

$$\lim_{M\to\infty}\limsup_{n\to\infty}\mathbb{E}\left[\tilde{Z}_{n,M}\mathbf{1}_{\{\mathfrak{R}_1^n\le N/\sigma_n\}}\right]=0. \tag{4.72}$$

By Markov's inequality, we have

$$\mathbb{P}(\tilde{Z}_{n,M}>\epsilon)\le\epsilon^{-1}\mathbb{E}[\tilde{Z}_{n,M}\mathbf{1}_{\{\mathfrak{R}_1^n\le N/\sigma_n\}}]+\mathbb{P}(\mathfrak{R}_1^n>N/\sigma_n).$$

According to (4.71) and (4.72), we can first choose some number $N=N(\epsilon)$ then some number $M=M(N(\epsilon),\epsilon)=M(\epsilon)$ such that $\limsup_n\mathbb{P}(\tilde{Z}_{n,M}>\epsilon)<\epsilon$. On the other hand, Corollary 3 of [41] says that $\Sigma(n,1,M)$ is distributed like $\tilde{Z}_{n,M}$. Then we have shown (4.62) for $k=1$. The general case can be treated in the same way, and we omit the details. □

So far we have completed the proof of Proposition 4.23 in the case where $\boldsymbol{\theta}$ has all strictly positive entries. The other cases are even simpler:

Case 2. Suppose that $\theta_0=0$, we take $m_n=n$ and the same argument follows.

Case 3. Suppose that $\boldsymbol{\theta}$ has a finite length I, then it suffices to take $m_n=I$. We can proceed as before.

Chapter 5

Reversing the cut tree of the Brownian continuum random tree

The results of this chapter are from the joint work [40] with Nicolas Broutin, submitted for publication.

Contents

We consider the logging process of the Brownian continuum random tree (CRT) $\mathcal{T}$ using a Poisson point process of cuts on its skeleton, as introduced in Aldous and Pitman [11]. The cut tree defined by Bertoin and Miermont [30] describes the genealogy of the fragmentation of $\mathcal{T}$ into connected components. This cut tree $\mathrm{cut}(\mathcal{T})$ is distributed as another Brownian CRT, and is a function of the original tree $\mathcal{T}$ and of the randomness in the logging process. We are interested in reversing the transformation of $\mathcal{T}$ into $\mathrm{cut}(\mathcal{T})$: we define a *shuffling* operation, which given a Brownian CRT $\mathcal{H}$, yields another one $\mathrm{shuff}(\mathcal{H})$ distributed in such a way that $(\mathcal{T}, \mathrm{cut}(\mathcal{T}))$ and $(\mathrm{shuff}(\mathcal{H}), \mathcal{H})$ have the same distribution.

5.1 Introduction

Let $\mathcal{T}$ be Aldous' Brownian continuum random tree (CRT) [8]. To the logging process of $\mathcal{T}$ introduced in Aldous and Pitman [13], one can associate another continuum random tree $\mathrm{cut}(\mathcal{T})$, which describes the genealogical structure of this fragmentation process (see [30] and [39]). Moreover, for a Brownian CRT, this associated tree $\mathrm{cut}(\mathcal{T})$ is also distributed as a Brownian CRT. One of the main questions [Miermont,

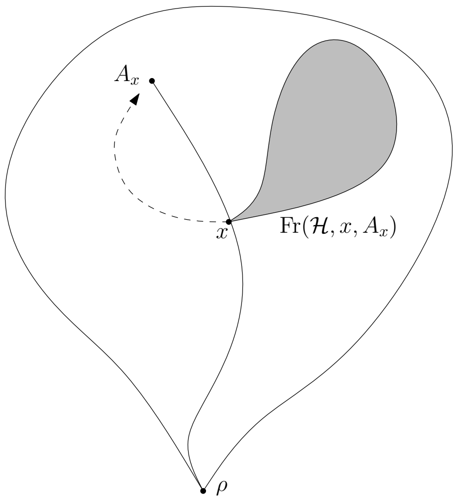

Figure 5.1 – *The surgical operation on the tree $\mathcal{H}$, rooted at ρ, for a single branch point x with corresponding attach point A_x.*

Pers. Comm.] then is whether the transformation from $\mathcal{T}$ to the genealogy of the fragmentation $\mathrm{cut}(\mathcal{T})$ is "reversible". Of course, some information has been lost about the initial tree $\mathcal{T}$, and one must first understand whether it is possible to resample this information, and then study the possibility of the construction of a tree $\mathcal{T}'$ that is distributed like $\mathcal{T}$, conditional on $\mathrm{cut}(\mathcal{T})$.

Let $\mathcal{H}$ be another Brownian CRT (that should be informally thought of as $\mathrm{cut}(\mathcal{T})$), we define below a continuous tree $\mathrm{shuff}(\mathcal{H})$, which is random given $\mathcal{H}$, and such that the following identity in distribution holds:

$$\big(\mathrm{shuff}(\mathcal{H}), \mathcal{H}\big) \stackrel{d}{=} \big(\mathcal{T}, \mathrm{cut}(\mathcal{T})\big). \tag{5.1}$$

The construction of the tree $\mathrm{shuff}(\mathcal{H})$ from $\mathcal{H}$ is the main objective of the present document, and can be seen as follows. Let $\mathrm{Br}(\mathcal{H})$ denote the set of branch points of $\mathcal{H}$. Start by assigning independently to every branch point $x \in \mathrm{Br}(\mathcal{H})$ a random point A_x sampled using the mass measure ν restricted to $\mathrm{Sub}(\mathcal{H}, x)$, the subtree of $\mathcal{H}$ above x. For each such x, the choice of A_x induces a choice among the subtrees of $\mathrm{Sub}(\mathcal{H}, x)$ rooted at x: Let the *fringe* $\mathrm{Fr}(\mathcal{H}, x, A_x)$ be the subset of points $y \in \mathrm{Sub}(\mathcal{H}, x)$ for which the closest common ancestor of y and A_x is $y \wedge A_x = x$. Then, informally $\mathrm{shuff}(\mathcal{H})$ is obtained by detaching $\mathrm{Fr}(\mathcal{H}, x, A_x)$ and reattaching it at A_x, for every branch point x of $\mathcal{H}$ (see Figure 5.1); the points of the skeleton that are not branch points are not used. It is a priori unclear whether this definition makes sense, let alone that the resulting metric space is a real tree or that it has the correct distribution. It indeed seems that we discard from $\mathcal{H}$ all the length by leaving the skeleton behind. The remainder of the document is devoted to making this construction rigorous, and to prove that the tree $\mathrm{shuff}(\mathcal{H})$ satisfies (5.1).

The continuous problem at hand is connected to a rather large body of work on the destruction of random (discrete) trees by sampling of random nodes or edges initiated by Meir and Moon [88]. There is no significant difference between sampling nodes or edges, and we present here a version that samples nodes and proceeds as follows: sample a random node in some rooted tree (random or not), discard the portion that is now disconnected from the root, and keep going until the root is finally picked (the process then stops). The main question addressed by Meir and Moon [88] and many of the researchers after them was about the number of steps, or cuts, that are needed for the process to terminate. This problem has been considered for a number of classical models of trees including random binary search trees [67, 68], random recursive trees [21, 28, 48, 70] and the family trees of Galton–Watson processes conditioned on the total progeny [7, 60, 72, 93]. Janson [72] was the first to realize by moment calculations that, when the tree is a Galton–Watson tree, there should be nice constructions of the limit random variables directly in terms of continuous cutting of the Brownian CRT. In some sense, the continuous cutting alluded in [72] is just a version of the logging of the CRT in which only the cuts affecting the size of the connected component containing the root are retained. The constructions in [7], [27] and [30] all encode the "number of cuts" affecting the connected components containing some points as the total length of some distinguished subtrees. The construction of the genealogy as a compact tree is due to Bertoin and Miermont [30].

Let us now describe our approach to the definition of $\mathrm{shuff}(\mathcal{H})$. The idea is to construct it by defining an order in which the fringes should be sent to their new attach points. This ordering yields a tree-valued Markov chain, and we formally define $\mathrm{shuff}(\mathcal{H})$ as its almost sure limit. More precisely, a construction of the first element of this Markov chain appears in [7] and has been formally justified in [39]: there the subtrees to be reattached are only those lying along the path between the root and a distinguished random leaf U_1. In the following, we refer to this transformation as the *one-path reversal*. The Markov chain we have in mind consists in iteratively reattaching the subtrees lying on the paths to an i.i.d. sequence of leaves $(U_i)_{i\geq 1}$ in $\mathcal{H}$, that we later refer to as the *i-paths reversals*, or *i-reversals* for short. However, not any such sequence $(U_i)_{i\geq 1}$ would do. Indeed, although it is very close to the one we are after, the one-path reversal enforces that the subtrees to be detached are precisely $\mathrm{Fr}(\mathcal{H}, x, U_1)$, for the branch points x on the path to U_1; so in particular, they are only defined in terms of U_1, and the choices A_x are then somewhat conditioned on being consistent with the constraint that

$$\mathrm{Fr}(\mathcal{H}, x, A_x) = \mathrm{Fr}(\mathcal{H}, x, U_1). \tag{5.2}$$

It follows that if we want to use the results of [7, 39], then the sequence $(U_i)_{i\geq 1}$ must be constructed from $(A_x, x \in \mathrm{Br}(\mathcal{H}))$ in such a way that, for all the branch points x on the path to U_i, the constraint in (5.2) is satisfied with U_i instead of U_1 in the right-hand side.

Plan of the chapter. The route we use here to define $\mathrm{shuff}(\mathcal{H})$ relies on a careful understanding of the cutting procedure and of the genealogy induced by finitely many random points only. In Section 5.2, we introduce the relevant background on cut trees, and their reversals. We also prove a few results that have not appeared elsewhere. Section 5.3 is devoted to proving that the sequence of k-paths reversals converges as $k \to \infty$ in the sense of Gromov–Prokhorov. Up to this point, the shuffle operation is therefore justified as a refining sequence of k-reversals. The direct construction presented above is then justified in Section 5.4 by proving that one can construct a sequence of leaves such that the shuffle tree corresponds to the limit of the k-reversals with respect to this sequence of leaves. Some auxiliary results about the Brownian CRT for which we did not find a reference are proved in Section 5.5.

5.2 Preliminaries on cut trees and shuffle trees

In this section, we recall the previous results in [39] on the cut trees and the shuffle trees of the Brownian continuum random tree.

5.2.1 Notations and background on continuum random trees

We only give here a short overview, the interested reader may consult [10], [82], or [57] for more details.

A *real tree* is a geodesic metric space without loops. The real trees we are interested in are compact. A *continuum random tree* $\mathcal{T}$ is a random (rooted) real tree equipped with a probability measure, often referred to as the *mass measure* or the uniform measure. The *Brownian continuum random tree* is a special continuum random tree that has been introduced by Aldous [8] as the scaling limit of uniformly random trees. One way to define the Brownian CRT starts from a standard normalized Brownian excursion of unit length $e = (e_s, 0 \leq s \leq 1)$. For any $s, t \in [0, 1]$, let

$$\frac{1}{2}d(s,t) := e_s + e_t - 2 \inf_{u \in [s,t]} e_u, \tag{5.3}$$

and define $s \sim t$ if $d(s,t) = 0$. Then d induces a metric on the quotient space $[0,1]/\sim$. Moreover, this metric space is a real tree: it is the Brownian CRT, which we denote by $(\mathcal{T}, d_{\mathcal{T}})$ in the following. A Brownian CRT $\mathcal{T}$ also comes with a *mass measure* $\mu_{\mathcal{T}}$, which is the push-forward of Lebesgue measure by the canonical projection $p : [0,1] \to \mathcal{T}$. A point that is sampled according to $\mu_{\mathcal{T}}$ is usually called here a $\mu_{\mathcal{T}}$-point. The Brownian CRT $\mathcal{T}$ is rooted at the point $p(0)$. For $a > 0$, let $e^{(a)} = (e_s^{(a)}, 0 \leq s \leq a)$ denote the Brownian excursion of length a. We can associate with $e^{(a)}$ a random real tree, denoted by $\mathcal{T}^{(a)}$, by replacing e with $e^{(a)}$ in (5.3). If $s > 0$, we denote by $s\mathcal{T}$ the metric space in which the distance is $sd_{\mathcal{T}}$. Then, the Brownian scaling implies that (see also Section 5.5)

$$\mathcal{T}^{(a)} \stackrel{d}{=} \sqrt{a}\mathcal{T}.$$

And clearly, the mass measure of $\mathcal{T}^{(a)}$, which is the push-forward of Lebesgue measure on $[0,a]$, has total mass a. In what follows, we sometimes refer to $\mathcal{T}$ as the standard Brownian CRT.

For $u, v \in \mathcal{T}$, we denote by $[\![u, v]\!]$ and $]\!]u, v[\![$ the closed and open paths between u and v in $\mathcal{T}$, respectively. For $u \in \mathcal{T}$, the degree of u in $\mathcal{T}$, denoted by $\deg(u, \mathcal{T})$, is the number of connected components of $\mathcal{T} \setminus \{u\}$. We also denote by

$$\mathrm{Lf}(\mathcal{T}) = \{u \in \mathcal{T} : \deg(u, \mathcal{T}) = 1\} \quad \text{and} \quad \mathrm{Br}(\mathcal{T}) = \{u \in \mathcal{H} : \deg(u, \mathcal{T}) \geq 3\}$$

the set of the *leaves* and the set of *branch points* of $\mathcal{T}$, respectively. Almost surely, these two sets are everywhere dense in $\mathcal{T}$, though $\mathrm{Lf}(\mathcal{T})$ is uncountable and $\mathrm{Br}(\mathcal{T})$ countable. The skeleton of $\mathcal{T}$ is the complement of $\mathrm{Lf}(\mathcal{T})$ in $\mathcal{T}$, denoted by $\mathrm{Sk}(\mathcal{T})$. The skeleton is the union

$$\mathrm{Sk}(\mathcal{T}) = \bigcup_{u,v \in \mathcal{T}}]\!]u, v[\![.$$

If ρ is the root of $\mathcal{T}$, for $u \in \mathcal{T}$, the subtree above u, denoted by $\mathrm{Sub}(\mathcal{T}, u)$, is defined to be the subset $\{v \in \mathcal{T} : u \in [\![\rho, v]\!]\}$. If $v \in \mathrm{Sub}(\mathcal{T}, u)$ is distinct from u, we denote by $\mathrm{Fr}(\mathcal{T}, u, v)$ the fringe tree hung from $[\![u, v]\!]$ which is the set $\{w \in \mathrm{Sub}(\mathcal{T}, u) : [\![w, u]\!] \cap [\![u, v]\!] = \{u\}\}$. It is nontrivial only if $u \in \mathrm{Br}(\mathcal{T})$. There also exists a unique σ-finite measure ℓ concentrated on $\mathrm{Sk}(\mathcal{T})$ such that for any two points $u, v \in \mathcal{T}$ we have $\ell([\![u, v]\!]) = d_{\mathrm{T}}(u, v)$; ℓ is called the *length measure*. If ρ denotes the root of $\mathcal{T}$ and $v_1, \cdots, v_k$ are k points of $\mathcal{T}$, we write

$$\mathrm{Span}(\mathrm{T}; v_1, \cdots, v_k) = \bigcup_{1 \leq i \leq k} [\![\rho, v_i]\!]$$

for the subtree of $\mathcal{T}$ spanning the root ρ and $\{v_i, 1 \leq i \leq k\}$.

The state space of interest is the set of metric spaces that are pointed, that is with a distinguished point that we call the root and equipped with a probability measure. More precisely, it is the set of equivalence classes induced by measure-preserving isometries (on the support of the probability measure). When equipped with the Gromov–Prokhorov (GP) distance, this yields a Polish space. Convergence in the GP topology is equivalent to convergence in distribution of the matrices whose entries are distances between the pairs of points sampled from the probability distribution μ. This is discussed at length in [39], and we also refer the reader to [64] and [62] for more information.

5.2.2 The cutting procedure on a Brownian CRT

Let $\mathcal{T}$ be a Brownian CRT to be cut down. Now let $\mathcal{P}$ be a Poisson point process of intensity measure $dt \otimes \ell(dx)$ on $\mathbb{R}_+ \times \mathcal{T}$. Every point $(t, x) \in \mathcal{P}$ is seen as a cut on $\mathcal{T}$ at location x which arrives at time t. Then $\mathcal{P}$ defines a Poisson rain of cuts that split $\mathcal{T}$ into smaller and smaller connected components as time goes. More precisely, let $(V_i)_{i\geq 1}$ be a sequence of independent points sampled according to μ, then for each $t \geq 0$, $\mathcal{P}$ induces a nested process of exchangeable partitions of $\mathbb{N}$ in the following way. For each $t \geq 0$, the blocks are the equivalence classes of the relation $\sim_t$ defined by

$$i \sim_t j \quad \text{if and only if} \quad \mathcal{P} \cap ([0,t] \times [\![V_i, V_j]\!]) = \varnothing.$$

Let $\mathcal{T}_i(t)$ be the set of those points in $\mathcal{T}$ which are still connected to V_i at time t, that is

$$\mathcal{T}_i(t) := \{u \in \mathcal{T} : \mathcal{P} \cap ([0,t] \times [\![V_i, u]\!]) = \varnothing\}.$$

Then it is easy to see that $\mathcal{T}_i(t)$ is a connected subspace of $\mathcal{T}$, that is, a subtree of $\mathcal{T}$. Furthermore, we have $\mathcal{T}_i(t) \subseteq \mathcal{T}_i(s)$ if $s \leq t$, and $\cap_{t\geq 0}\mathcal{T}_i(t) = \{V_i\}$ almost surely, since with probability one the atoms of $\mathcal{P}$ are everywhere dense in $\mathcal{T}$ and V_i is not among these atoms.

5.2.3 The k-cut tree

The main point of the definition of a cut tree is to obtain a representation of the genealogy of the fragmentation induced by $\mathcal{P}$ as a compact real tree. A first step consists in focusing on the genealogy of the fragmentation induced on $[k] = \{1, 2, \ldots, k\}$, for some $k \geq 1$. So at this point, we only keep track of the evolution of the connected components containing the points $V_1, V_2, \cdots, V_k$ and ignore all the other ones.

For each $t \geq 0$, let us write $\pi_k(t)$ for the partition of $[k]$ induced by $\sim_t$. Then for any $t \geq s$, $\pi_k(t)$ is a refinement of $\pi_k(s)$. We encode the family $(\pi_k(t))_{t\geq 0}$ by a rooted tree S_k with k leaves. Each equivalence class induced on $[k]$ by some $\sim_t$, $t \geq 0$ is represented by a *node* of S_k. It is also convenient to add an additional node r, which we see as the root of S_k. From the root r there is a unique edge, which connects r to the node labeled by $[k] := \{1, 2, \cdots, k\}$. Let $t_{[k]} := \sup\{t \geq 0 : \pi_k(t) \neq \{[k]\}\}$ be the time when $[k]$ disappears from $(\pi_k(t))_{t\geq 0}$. Note that $t_{[k]}$ is the first moment when there is some point $x_{[k]}$ of the subtree of $\mathcal{T}$ spanning $\{V_1, V_2, \ldots, V_k\}$,

$$\bigcup_{1\leq i,j\leq k} [\![V_i, V_j]\!]$$

such that $(t_{[k]}, x_{[k]}) \in \mathcal{P}$. Observe that almost surely $x_{[k]}$ has degree two in $\mathcal{T}$, so that $\pi_k(t_{[k]})$ consists of only two blocks E_1 and E_2 with probability one. This is represented in S_k by the fact that the node labelled $[k]$ has two children, labeled respectively by E_1 and E_2. One then proceeds recursively to define the subtrees induced on the set of leaves in E_1 and E_2, respectively. We obtain a binary tree on k leaves labelled by $\{1\}, \cdots, \{k\}$. (See Figure 5.2).

We now endow S_k with a distance d_{S_k}, or to be more precise we define a binary real tree that has the same tree structure as S_k. For this, we set for $1 \leq i \leq k$ and $t \in [0, \infty]$,

$$L_i(t) := \int_0^t \mu\big(\mathcal{T}_i(s)\big) ds.$$

Then $L_i(\infty)$ is finite almost surely (this is shown for instance in [7]). For every $i \in [k]$, we want to identify the unique path of S_k from the root $[k]$ to the leaf $\{i\}$ with the finite interval $\{L_i(t) : t \in [0, \infty]\}$ such that if a node is labelled by E for some $E \subseteq [k]$, then it is at distance $L_i(t_E)$ from the root,

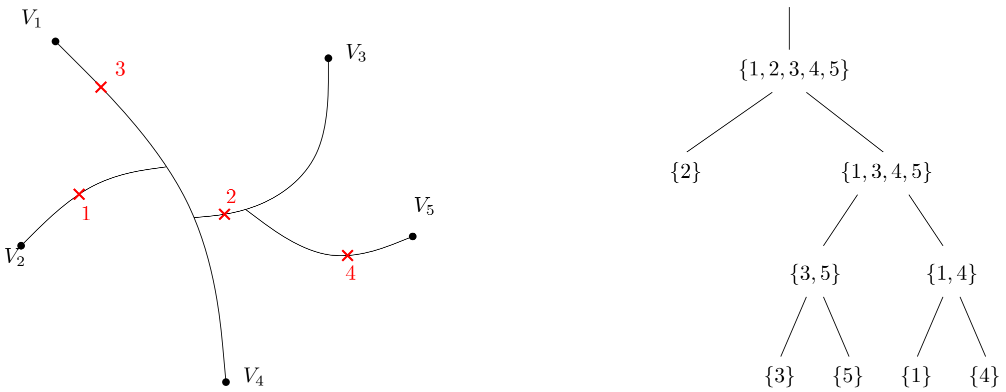

Figure 5.2 – *On the left, the subtree of $\mathcal{T}$ spanning the leaves $V_1, V_2, \ldots, V_5$. The cuts falling on it are represented by the crosses, and the index next to them is the time at which they appear. On the right, the corresponding tree S_k.*

where $t_E = \sup\{t \geq 0 : E \in \pi_k(t)\}$. Doing so does not cause any ambiguity since if $i, j \in E$ then $\mathcal{T}_j(s) = \mathcal{T}_i(s)$ for any $s \leq t_E$. So we obtain a compact real tree which consists of k paths of respective lengths $L_i(\infty), 1 \leq i \leq k$. By slightly abusing the notation, we still write S_k for the real tree (S_k, d_{S_k}). In other words, if we write $E_i(0) = r, E_i(1) = [k], E_i(2), \ldots, E_i(h_i) = \{i\}$ for the sequence of nodes on the path from the root to $\{i\}$ in S_k, the real tree (S_k, d_{S_k}) is the tree S_k in which the edges have been replaced by the $2k-1$ intervals of lengths $L_i(t_{E_i(h+1)}) - L_i(t_{E_i(h)})$, for $0 \leq h < h_i$ and $1 \leq i \leq k$.

We now move on to the definition of the k-cut tree. The real tree S_k provides the *backbone* of the k-cut tree. We define the k-cut tree $\mathrm{cut}(\mathcal{T}, V_1, \cdots, V_k)$ as the real tree obtained by grafting on the backbone S_k the subtrees discarded during the cutting procedure. Let $\mathcal{C}^k$ be the set of those $t \geq 0$ for which

$$\mu\Big(\bigcup_{1\leq i\leq k} \mathcal{T}_i(t)\Big) < \mu\Big(\bigcup_{1\leq i\leq k} \mathcal{T}_i(t-)\Big),$$

Note that $\mathcal{C}^k$ is almost surely countable. If $t \in \mathcal{C}^k$, let i_t be the smallest element of $[k]$ such that $\Delta\mathcal{T}_{i_t}(t) := \mathcal{T}_{i_t}(t-) \setminus \bigcup_{1\leq j\leq k} \mathcal{T}_j(t) \neq \varnothing$. Because of the holes left by the previous cuts, $\Delta\mathcal{T}_{i_t}(t)$ is a connected but not complete subspace in $\mathcal{T}$. We let Δ_t^k be the completion of $\Delta\mathcal{T}_{i_t}(t)$. Almost surely there exists a unique $x \in \mathcal{T}$ such that $(t, x) \in \mathcal{P}$. Note that $x \in \Delta\mathcal{T}_{i_t}(t)$. We denote by $x' \in \Delta_t^k$ the image of x via the canonical injection from $\Delta\mathcal{T}_{i_t}(t)$ to Δ_t^k. We think of Δ_t^k as rooted at x'. Then for each $t \in \mathcal{C}^k$, we graft Δ_t^k by its root on the path in S_k connecting the root to the leaf $\{i_t\}$ at distance $L_{i_t}(t)$ from the root. We denote by $\mathcal{G}_k = \mathrm{cut}(\mathcal{T}, V_1, \cdots, V_k)$ the obtained metric space. The tree $\mathcal{G}_k$ also bears a mass measure which is inherited from that of $\mathcal{T}$; the set of the points which have been added (either in the backbone S_k or due to completion) is assigned mass 0. The new mass is still denoted by μ. An alternative way to define $\mathcal{G}_k$ (which is the way we have used in [39]) is to graft $\Delta\mathcal{T}_{i_t}(t)$ (rather than Δ_t^k) on S_k, and then to complete the metric space. One easily checks that these two definitions coincide.

Remark. There is a number of different mass measures that we need to consider here. In order to clarify the discussion and to keep the notation under control, we have decided to keep using the same name for the mass measure when only a set of measure zero was modified by the transformation either by removal of countably many points, by (countable) completion, or by the addition of a backbone. For instance, we think of the tree $\mathcal{G}_k$ as still carrying the mass measure μ of $\mathcal{T}$.

Proposition 5.1 (Distribution of the k-cut tree)**.** *If $\mathcal{T}$ is the Brownian CRT, and $(V_i)_{i\geq 1}$ is a sequence of i.i.d. points of $\mathcal{T}$ with common distribution μ, then for each $k \geq 1$, we have*

$$(\mathcal{G}_k, S_k) \stackrel{d}{=} \big(\mathcal{T}, \mathrm{Span}(\mathcal{T}; V_1, V_2, \cdots, V_k)\big). \tag{5.4}$$

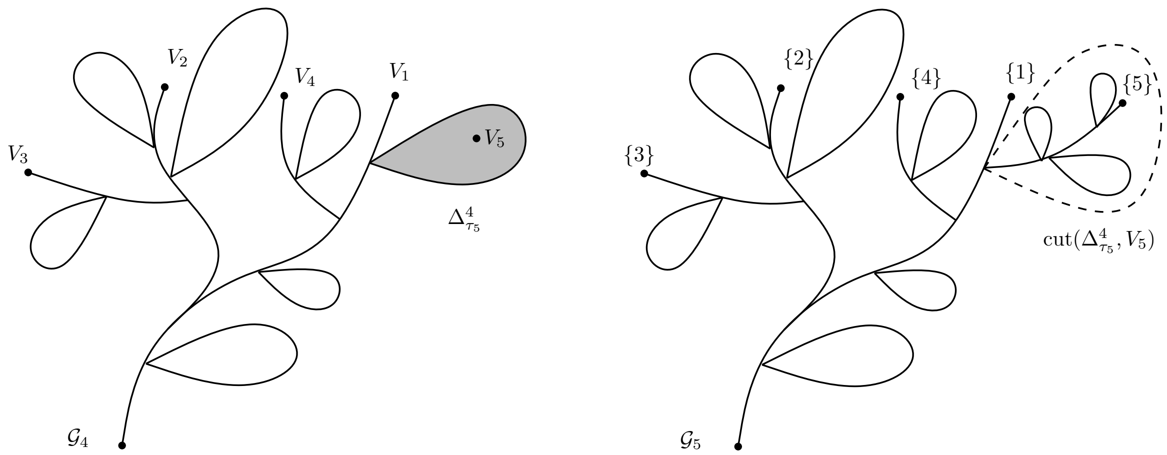

Figure 5.3 – *The transformation from* $\mathcal{G}_{k-1} = \mathrm{cut}(\mathcal{T}, V_1, \ldots, V_{k-1})$ *to* $\mathcal{G}_k = \mathrm{cut}(\mathcal{T}, V_1, \ldots, V_{k-1}, V_k)$ *with* $k = 5$*: with probability one,* V_k *falls in some* $\Delta^{k-1}_{\tau_k}$*, a connected component of* $\mathcal{G}_{k-1} \setminus S_{k-1}$*. That subtree stopped being transformed at time* τ_k*, since it did not contain any of* $V_1, \ldots, V_{k-1}$*, and it should now be replaced by* $\mathrm{cut}(\Delta^{k-1}_{\tau_k}; V_k)$*.*

The case $k = 1$ corresponds to a special case of Theorem 3.2 in [39]. The general case follows from the analogous result on the discrete trees (see Section 4, Lemma 4.5 there) and the same weak convergence argument as in [39], and we omit the details.

The "complete" cut tree By construction, since the definition of $\mathcal{T}_i(t)$ does not depend on k, we have $S_k \subset S_{k+1}$ for each $k \geq 1$.

Proposition 5.2 (Complete cut tree, [30, 39])**.** *Let* $\mathrm{cut}(\mathcal{T}) = \overline{\cup_k S_k}$ *be the limit metric space of* $(S_k)_{k\geq 1}$*. If* $\mathcal{T}$ *is the Brownian CRT, then almost surely,* $\mathrm{cut}(\mathcal{T})$ *is a compact real tree and is distributed as* $\mathcal{T}$*.*

The construction of $\mathcal{G}_k$ described above yields the following recurrence relation between $\mathcal{G}_{k+1}$ and $\mathcal{G}_k$ (see Figure 5.3). For every $k \geq 1$, the collection Δ^k_t, $t \in \mathcal{C}^k$ has full mass and a uniform point V falls with probability one in Δ^k_t, for some $t \in \mathcal{C}^k$. If we let $m_k := \mu(\Delta^k_t)$, then $\mathrm{m}_k{}^{-1/2}\Delta^k_t$ is distributed as a standard Brownian CRT. As a consequence, the 1-cut tree $\mathrm{cut}(\Delta^k_t, V)$ is well-defined.

Proposition 5.3 (Recurrence relation for $(\mathcal{G}_k)_{k\geq 0}$)**.** *For each* $k \geq 2$*, let* $\tau_k \in \mathcal{C}^{k-1}$ *be such that* $V_k \in \Delta^{k-1}_{\tau_k}$*. Then* $\mathcal{G}_k$ *is obtained from* $\mathcal{G}_{k-1}$ *by replacing* $\Delta^{k-1}_{\tau_k}$ *with* $\mathrm{cut}(\Delta^{k-1}_{\tau_k}, V_k)$*.*

The cut tree $\mathrm{cut}(\mathcal{T})$ may then be seen as the limit of the k-cut trees. The notion of $\mathrm{cut}(\mathcal{T})$ here coincides with that in [28].

Proposition 5.4 (Convergence of k-cut trees)**.** *Let* $\mathcal{T}$ *be the Brownian CRT. As* $k \to \infty$*,*

$$\mathrm{cut}(\mathcal{T}, V_1, \cdots, V_k) \to \mathrm{cut}(\mathcal{T}), \quad \textit{almost surely}$$

in the Gromov–Hausdorff topology.

Proof. If (T, d) is a compact real tree rooted at ρ, we let $\mathrm{ht}(T) = \sup_{u\in T} d(u, \rho)$ and $\mathrm{diam}(T) = \sup_{u,v\in T} d(u, v)$ denote the respective height and diameter of T. By the triangle inequality, we have $\mathrm{ht}(T) \leq \mathrm{diam}(T) \leq 2\,\mathrm{ht}(T)$. On the one hand, we deduce from Proposition 5.3 that

$$\Upsilon_k := \sup_{t\in\mathcal{C}^k} \mathrm{diam}(\Delta^k_t), \quad k \geq 1$$

is a non-increasing sequence, so it converges almost surely to some random variable that we denote by Υ_∞. On the other hand, if we write δ_{H} for the Hausdorff distance (on the compact subsets of $\mathcal{G}_k$), then it follows from the construction of $\mathcal{G}_k$ that

$$\delta_{\mathrm{H}}(\mathcal{G}_k, S_k) = \sup_{t\in\mathcal{C}^k} \mathrm{ht}(\Delta_t^k).$$

Proposition 5.1 implies that the sequence $(\delta_{\mathrm{H}}(\mathcal{G}_k, S_k))_{k\geq 1}$ converges to 0 in distribution. Combining this with the bounds $\delta_{\mathrm{H}}(\mathcal{G}_k, S_k) \leq \Upsilon_k \leq 2\,\delta_{\mathrm{H}}(\mathcal{G}_k, S_k)$, we obtain that $\Upsilon_\infty = 0$, which then entails that $\delta_{\mathrm{H}}(\mathcal{G}_k, S_k) \to 0$ almost surely as $k\to\infty$. Since, by Proposition 5.2, we have $\lim_{k\to\infty}\delta_{\mathrm{H}}(\mathrm{cut}(\mathcal{T}), S_k) = 0$ a.s., the result follows. $\square$

Distribution of a uniform path One of the main ingredients of our construction of the shuffle tree in Section 5.3 consists in understanding how the path between two points gets transformed as the approximation $\mathcal{G}_k$ of the cut tree $\mathrm{cut}(\mathcal{T})$ gets refined. More precisely, let ξ_1, ξ_2 be two independent μ-points of $\mathcal{T}$ and write $p :=]\!]\xi_1, \xi_2[\![$ for the open path between them in $\mathcal{T}$. Initially, in $\mathcal{G}_0 := \mathcal{T}$, p is indeed a path. But later on, as k increases, this path p gets cut into pieces each contained in some of the Δ_t^k that are grafted onto the backbone S_k. More formally, for each $k\geq 0$, there exists an injective map $\phi_k^\circ : \cup_{t\in\mathcal{C}^k}\Delta_t^k \to \mathcal{G}_k$ whose restriction to every $\mathrm{Sk}(\Delta_t^k)$, $t\in\mathcal{C}^k$, is the identity and that is continuous on Δ_t^k. Then with probability one, almost every point of p is contained in $\cup_{t\in\mathcal{C}^k}\Delta_t^k$; only the cut points are lost. To deal with this, let σ_k° denote the map acting on subsets of $\mathcal{T}$ which removes the points x such that $(t,x)\in\mathcal{P}$ and $t\in\mathcal{C}^k$; so the image of σ_k° is contained in $\cup_{t\in\mathcal{C}^k}\Delta_t^k$. Then, we set $\phi_k := \phi_k^\circ\circ\sigma_k^\circ$ and $p_k := \phi_k(p)$ is a union of disjoint open paths $\mathcal{G}_k$.

By the recursive construction in Proposition 5.3, understanding how the path p gets mapped into $\mathcal{G}_k$ by ϕ_k reduces to understanding one step of the transformation, that is how p gets mapped into $\mathcal{G}_1 = \mathrm{cut}(\mathcal{T}, V_1)$ by ϕ_1. The following result for $k=1$ has been proved in [39]. It is the basis of the one-path reversal of the next section, and is used in Section 5.3.5 to derive the distribution of $p_k = \phi_k(p)$.

Proposition 5.5 (Distribution of p_1, [39]). *Almost surely, there exist $M_1, M_2 \geq 0$ and two finite sequences of elements of $\mathcal{C}^1$:*

$$0 < t_{1,0} < t_{1,1} < \cdots < t_{1,M_1} \qquad \textit{and} \qquad 0 < t_{2,0} < t_{2,1} < \cdots < t_{2,M_2},$$

which are all distinct except that $t_{1,M_1} = t_{2,M_2}$, and there exist two sequences $(a_1'(m))_{0\leq m\leq M_1}$ and $(a_2'(m))_{0\leq m\leq M_2}$ satisfying $a_i'(0) = \xi_i$ and $a_i'(m)\in\Delta^1_{t_{i,m}}$ for $0\leq m\leq M_i$ such that

$$p_1 = \phi_1(p) = \bigcup_{m=0}^{M_1-1}]\!]a_1'(m), x_1'(m)[\![\cup \bigcup_{m=0}^{M_2-1}]\!]a_2'(m), x_2'(m)[\![\cup]\!]a_1'(M_1), a_2'(M_2)[\![. \tag{5.5}$$

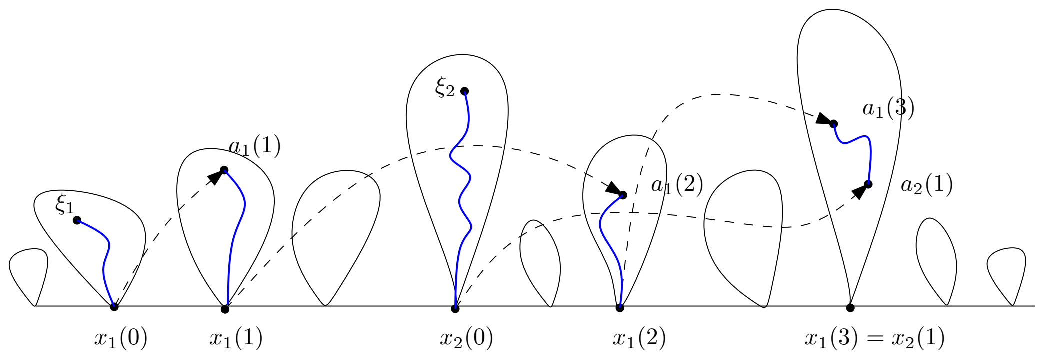

Figure 5.4 – *An example with $\mathscr{I}(1,2) = 3$, $\mathscr{I}(2,1) = 1$ and $\mathrm{mg}(1,2) = 4$. The dashed lines indicate the point identifications for the root of the relevant subtrees.*

Here, $x_i'(m)$ denotes the root of $\Delta^1_{t_{i,m}}$, for $0 \le m < M_i$, $i = 1, 2$. Moreover, conditionally on $\mathrm{cut}(\mathcal{T}, V_1)$*, we have the following description of their distributions: let $(a_i''(m))_{m \ge 0}$, $i = 1, 2$, be two independent Markov chains with the following distribution: $a_i''(0)$ is a μ-point; given $t_i''(m)$, $a_i''(m+1)$ is distributed according to the restriction of μ to $\cup_{t > t_{i,m-1}} \Delta^1_t$ and $t_i''(m+1)$ is the element $t \in \mathcal{C}^1$ such that $a_i''(m+1) \in \Delta^1_t$. Moreover, if we let (M_1'', M_2'') be the pair of smallest integers (m_1, m_2) satisfying $t_1''(m_1) = t_2''(m_2)$, then $M_i'' < \infty$ almost surely, $i = 1, 2$, and we have*

$$(a_i'(m))_{0 \le m \le M_i} \stackrel{d}{=} (a_i''(m))_{0 \le m \le M_i''}.$$

5.2.4 One-path reverse transformation and the 1-shuffle tree

As we have just seen, the cut tree $\mathcal{G}_1 = \mathrm{cut}(\mathcal{T}, V_1)$ has a distinguished path along which were grafted the subtrees that were pruned. Quite intuitively, the reversal should consist in taking that distinguished path, removing it, and reattaching the connected components thereby created back where they used to be in order to make a new tree. One should keep the intuition of Proposition 5.5 in mind; however, we do not want to assume that the reversal be performed on a tree that can indeed be obtained as $\mathrm{cut}(\mathcal{T}, V_1)$ for some $\mathcal{T}$ and V_1 and the results are stated in distribution.

Let $\mathcal{H}$ be a Brownian CRT rooted at ρ, and U_1 a point of $\mathcal{H}$ sampled according to the mass measure ν. By Proposition 5.1, $\mathcal{H}$ has the same distribution as $\mathrm{cut}(\mathcal{T}, V_1)$, where $\mathcal{T}$ is a Brownian CRT and V_1 is sampled in $\mathcal{T}$ according to the mass measure. Furthermore, we have constructed in [39] a real tree $\mathcal{Q}_1 = \mathrm{shuff}(\mathcal{H}, U_1)$ such that the following identity in distribution holds:

$$\big(\mathrm{shuff}(\mathcal{H}, U_1), \mathcal{H}\big) \stackrel{d}{=} \big(\mathcal{T}, \mathrm{cut}(\mathcal{T}, V_1)\big). \tag{5.6}$$

In particular, (5.6) implies that $\mathcal{Q}_1$ and $\mathcal{T}$ have the same distribution.

Let us now recall briefly the construction of $\mathcal{Q}_1$. Let $\{F_i^\circ\}_{i \ge 1}$ be the collection of connected components of $\mathcal{H} \setminus [\![\rho, U_1]\!]$ which have positive ν-mass (this is indeed countable). For each F_i°, there exists a unique point $x \in [\![\rho, U_1]\!]$ such that $F_x := F_i^\circ \cup \{x\}$ is connected in $\mathcal{H}$. Let $\mathbf{B}$ be the set of those points x. For two points x, x' of $\mathbf{B}$, we write $x \succ x'$ if $d_{\mathcal{H}}(\rho, x) > d_{\mathcal{H}}(\rho, x')$. Then for each $x \in \mathbf{B}$, we associate an attaching point A_x, which is independent and sampled according to the restriction of ν to $\cup_{x' \succ x} F_{x'}$.

By Proposition 5.5, in a space where we would have $\mathcal{H} = \mathrm{cut}(\mathcal{T}, V_1)$ and where $[\![\rho, U_1]\!]$ would be the distinguished path created, it would be possible to couple these choices with the cutting procedure in such a way that each A_x corresponds to the location in the initial tree $\mathcal{T}$ where F_x was detached from. Informally, if we were to glue these F_x back at A_x, we should obtain $\mathcal{T}$ back. So these choices are the correct ones, but nevertheless this transformation is *a priori* not well-defined, contrary to the discrete one (see [39, Section 3]).

The formal justification of this reverse transformation requires first to verify that the distance between two independent ν-points is a.s. well-defined (here only finitely many reattaching operations are needed), and then to construct $\mathrm{shuff}(\mathcal{H}, U_1)$ as the continuum random tree corresponding to the matrix of distances between a sequence of i.i.d. ν-points. In other words, unlike $\mathrm{cut}(\mathcal{T}, V_1)$, we do not construct $\mathrm{shuff}(\mathcal{H}, U_1)$ by actually reassembling pieces of $\mathcal{H}$, but we construct a tree that has the same metric properties.

Proposition 5.6 (Construction of $\mathrm{shuff}(\mathcal{H}, U_1)$, [39])**.** *Let $(\eta_i)_{i \ge 1}$ be a sequence of independent points of $\mathcal{H}$ sampled according to the mass measure ν. For each $i \ge 1$, let $(a_i(m))_{m \ge 0}$ be a sequence of points obtained as follows: $a_i(0) = \eta_i$; inductively for $m \ge 1$, let $x_i(m-1)$ be the element of $\mathbf{B}$ such that $a_i(m-1) \in F_{x_i(m-1)}$, and set $a_i(m) = A_{x_i(m)}$. Then for each pair $i \ne j$, the following quantity is almost surely finite:*

$$\mathscr{I}(i, j) := \inf\{m \ge 0 : \exists\, m' \ge 0 \text{ such that } x_i(m) = x_j(m')\};$$

and if we let

$$\gamma(i,j) := \sum_{\ell=0}^{\mathscr{I}(i,j)-1} d_{\mathcal{H}}\big(a_i(\ell), x_i(\ell)\big) + \sum_{m=0}^{\mathscr{I}(j,i)-1} d_{\mathcal{H}}\big(a_j(m), x_j(m)\big) \\ + d_{\mathcal{H}}\big(a_i(\mathscr{I}(i,j)), a_j(\mathscr{I}(j,i))\big), \tag{5.7}$$

We think of the matrix $(\gamma(i,j))_{i,j\geq 1}$ *as the distance matrix between the points* $(\eta_i)_{i\geq 1}$ *in a new metric space. Then* $(\gamma(i,j))_{i,j\geq 1}$ *defines a CRT, which we root at* η_1 *and denote by* $\mathrm{shuff}(\mathcal{H}, U_1)$*. Moreover, we have*

$$\big(\mathrm{shuff}(\mathcal{H}, U_1), \mathcal{H}\big) \stackrel{d}{=} \big(\mathcal{T}, \mathrm{cut}(\mathcal{T}, V_1)\big). \tag{5.8}$$

Note that only the points $(\eta_i)_{i\geq 1}$ are kept from $\mathcal{H}$. The other ones are constructed by the metric space completion. Furthermore, the very construction of the tree $\mathcal{Q}_1 = \mathrm{shuff}(\mathcal{H}, U_1)$ implies that, if one denotes by ν_1 its mass measure, then $(\eta_i)_{i\geq 1}$ is a sequence of i.i.d. ν_1-points in $\mathcal{Q}_1$. Observe also that $\gamma(i,j)$ corresponds to the distance between η_i and η_j after grafting all the $F_{x_i(\ell)}$ at $a_i(\ell+1)$ for $\ell < \mathscr{I}(i,j)$ and all the $F_{x_j(m)}$ at $a_j(m+1)$ for $m < \mathscr{I}(j,i)$ (see Figure 5.4). Again, one may think of a coupling where the points η_i, $i \geq 1$, would be chosen by sampling $(\xi_i)_{i\geq 1}$ in $\mathcal{T}$. Almost surely, all these points are still in $\mathrm{cut}(\mathcal{T}, V_1)$.

5.2.5 Multiple-paths reversal and the k-shuffle tree

Once $\mathrm{shuff}(\mathcal{H}, U_1)$ has been properly defined, the k-shuffle tree $\mathcal{Q}_k = \mathrm{shuff}(\mathcal{H}, U_1, \cdots, U_k)$ is then defined by induction. Suppose that we have constructed $\mathcal{Q}_{k-1}$ from $\mathcal{H}$ for some $k \geq 2$. Let $\tilde{T}_k^\circ$ be the component of $\mathcal{H} \setminus \mathrm{Span}(\mathcal{H}; U_1, \cdots, U_{k-1})$ containing U_k, and let $\tilde{T}_k$ be the completion of $\tilde{T}_k^\circ$. If we write $\tilde{\mathrm{m}}_k := \mu(\tilde{T}_k)$, then $\tilde{\mathrm{m}}_k^{-1/2}\tilde{T}_k$ is distributed as a standard Brownian CRT and $\mathrm{shuff}(\tilde{T}_k, U_k)$ is thus well-defined by Proposition 5.6. Now let $\tilde{\mathcal{H}}_k$ be the tree obtained from $\mathcal{H}$ by replacing $\tilde{T}_k$ with $\mathrm{shuff}(\tilde{T}_k, U_k)$. Then we define $\mathcal{Q}_k := \mathrm{shuff}(\tilde{\mathcal{H}}_k, U_1, \cdots, U_{k-1})$. The following is a continuous analog of Proposition 4.8 in [39].

Proposition 5.7 (Distribution of the k-shuffle tree)**.** *For each* $k \geq 1$*, we have*

$$\big(\mathrm{shuff}(\mathcal{H}, U_1, \cdots, U_k), \mathcal{H}\big) \stackrel{d}{=} \big(\mathcal{T}, \mathrm{cut}(\mathcal{T}, V_1, \cdots, V_k)\big). \tag{5.9}$$

Proof. We proceed by induction on $k \geq 1$. The base case $k = 1$ is (5.8) from Proposition 5.6. Assume now that (5.9) holds for all natural numbers up to $k-1 \geq 1$. By the scaling property, $\tilde{\mathrm{m}}_k^{-1/2}\tilde{T}_k$, equipped with the restriction of ν to $\tilde{T}_k$, is distributed as $\mathcal{H}$, and is independent of $\mathcal{H} \setminus \tilde{T}_k$. Thus, we can apply the induction hypothesis to find that

$$\big(\mathrm{shuff}(\tilde{T}_k, U_k), \tilde{T}_k\big) \stackrel{d}{=} \big(\Delta^k_{\tau_k}, \mathrm{cut}(\Delta^k_{\tau_k}, V_k)\big), \tag{5.10}$$

as $\Delta^k_{\tau_k}$ has the same distribution as $\tilde{T}_k$. In particular, all four trees in (5.10) are Brownian CRTs and $\mathrm{shuff}(\tilde{T}_k, U_k)$ and $\tilde{T}_k$ have the same distribution, and we deduce from the definition of $\tilde{\mathcal{H}}_k$ that

$$\big(\tilde{\mathcal{H}}_k, \mathrm{shuff}(\tilde{T}_k, U_k)\big) \stackrel{d}{=} \big(\mathcal{H}, \tilde{T}_k\big). \tag{5.11}$$

It follows that $\tilde{\mathcal{H}}_k$ and $\mathcal{H}$ have the same distribution. Then by the induction hypothesis, we have

$$\big(\mathcal{Q}_k, \tilde{\mathcal{H}}_k\big) \stackrel{d}{=} \big(\mathcal{T}, \mathcal{G}_{k-1}\big). \tag{5.12}$$

Now $\Delta^k_{\tau_k}$ is the connected component of $\mathcal{G}_{k-1} \setminus S_{k-1}$ containing the leaf labelled as k. It is thus obtained in the same way as $\tilde{T}_k$ from $\mathcal{H}$. As a consequence, $(\mathcal{H} \setminus \tilde{T}_k, \tilde{T}_k)$ and $(\mathcal{G}_{k-1} \setminus \Delta^k_{\tau_k}, \Delta^k_{\tau_k})$ have the same distribution. Combining this with (5.11) and (5.12), we obtain

$$\big(\mathcal{Q}_k, \tilde{\mathcal{H}}_k \setminus \mathrm{shuff}(\tilde{T}_k, U_k), \mathrm{shuff}(\tilde{T}_k, U_k)\big) \overset{d}{=} \big(\mathcal{T}, \mathcal{G}_{k-1} \setminus \Delta^k_{\tau_k}, \Delta^k_{\tau_k}\big).$$

The transformation from $\tilde{T}_k$ to $\mathrm{shuff}(\tilde{T}_k, U_k)$ only involves sampling random points in $\tilde{T}_k$, and is therefore independent of the transformation from $\tilde{\mathcal{H}}_k$ to $\mathcal{Q}_k$. Similarly the transformations from $\mathcal{T}$ to $\mathcal{G}_{k-1}$ and the one from $\Delta^k_{\tau_k}$ to $\mathrm{cut}(\Delta^k_{\tau_k}, V_k)$ are also independent. Then it follows from (5.10) that

$$\big(\mathcal{Q}_k, \tilde{\mathcal{H}}_k \setminus \mathrm{shuff}(\tilde{T}_k, U_k), \mathrm{shuff}(\tilde{T}_k, U_k), \tilde{T}_k\big) \overset{d}{=} \big(\mathcal{T}, \mathcal{G}_{k-1} \setminus \Delta^k_{\tau_k}, \Delta^k_{\tau_k}, \mathrm{cut}(\Delta^k_{\tau_k}, V_k)\big)$$

Finally, observe that $\tilde{\mathcal{H}}_k \setminus \mathrm{shuff}(\tilde{T}_k, U_k) = \mathcal{H} \setminus \tilde{T}_k$ and that $\mathcal{G}_{k-1} \setminus \Delta^k_{\tau_k} = \mathcal{G}_k \setminus \mathrm{cut}(\Delta^k_{\tau_k}, V_k)$. This yields

$$\big(\mathcal{Q}_k, \mathcal{H} \setminus \tilde{T}_k, \tilde{T}_k\big) \overset{d}{=} \big(\mathcal{T}, \mathcal{G}_k \setminus \mathrm{cut}(\Delta^k_{\tau_k}, V_k), \mathrm{cut}(\Delta^k_{\tau_k}, V_k)\big),$$

which entails that (5.9) holds for k, and completes the proof. □

5.3 Convergence of k-shuffle trees and the shuffle tree

5.3.1 The shuffle tree

In this section, we prove the following result, which constitutes the foundations of the formal definition of the shuffle tree.

Theorem 5.8 (Convergence of the k-shuffle trees)**.** *For a.e. Brownian CRT $\mathcal{H}$, the limit of the sequence* $(\mathrm{shuff}(\mathcal{H}, U_1, \cdots, U_k))_{k\geq 1}$ *exists almost surely in the Gromov–Prokhorov topology.*

The sequence of leaves $(U_i)_{i\geq 1}$ that is used influences the limit: in particular, it determines which subtrees are fringes and in which direction they are sent to. Still, in the same way that $\mathrm{cut}(\mathcal{T})$ does depend on the cutting procedure, we denote the limit by $\mathrm{shuff}(\mathcal{T})$, although it does depend on the sequence $(U_i)_{i\geq 1}$. However, $(U_i)_{i\geq 1}$ *does not* contain all the randomness used in the construction. In Section 5.4, we construct a specific sequence of leaves which emphasizes the randomness hidden in the construction, justifying the claim in the introduction that the sequence $(A_x, x \in \mathrm{Br}(\mathcal{H}))$ is all that one needs.

The following is a direct consequence of Theorem 5.8, and justifies the claim that $\mathrm{shuff}(\,\cdot\,)$ is indeed the reverse transformation of $\mathrm{cut}(\,\cdot\,)$.

Corollary 5.9. *We have the following identity in distribution:*

$$\big(\mathrm{shuff}(\mathcal{H}), \mathcal{H}\big) \overset{d}{=} \big(\mathcal{T}, \mathrm{cut}(\mathcal{T})\big). \tag{5.13}$$

Proof. Let f and g be two bounded real-valued functions that are continuous in the Gromov–Prokhorov topology. Recall the notation $\mathcal{G}_k = \mathrm{cut}(\mathcal{T}, V_1, \cdots, V_k)$. By (5.9), we have

$$\mathbb{E}[f(\mathcal{Q}_k) \cdot g(\mathcal{H})] = \mathbb{E}[f(\mathcal{T}) \cdot g(\mathcal{G}_k)].$$

By Proposition 5.4 and the dominated convergence theorem, the right-hand side above converges to $\mathbb{E}[f(\mathcal{T}) \cdot g(\mathrm{cut}(\mathcal{T}))]$, as $k \to \infty$. Similarly, by Theorem 5.8, the left-hand side converges to $\mathbb{E}[f(\mathrm{shuff}(\mathcal{H})) \cdot g(\mathcal{H})]$. Therefore,

$$\mathbb{E}[f(\mathrm{shuff}(\mathcal{H}) \cdot g(\mathcal{H})] = \mathbb{E}[f(\mathcal{T}) \cdot g(\mathrm{cut}(\mathcal{T}))].$$

Since f and g were arbitrary, this entails (5.13). □

Let $(\eta_i)_{i\geq 1}$ be the sequence of independent points of $\mathcal{H}$ in Proposition 5.6, which is independent of the sequence $(U_i)_{i\geq 1}$. Recall the random variable $\gamma(i,j)$, which is the distance between η_i and η_j in $\mathcal{Q}_1$. Part of the statement of Proposition 5.6 says that $(\eta_i)_{i\geq 1}$ is a a family of i.i.d. uniform points in $\mathcal{Q}_1$. Because of the inductive definition of $\mathcal{Q}_k$, the sequence $(\eta_i)_{i\geq 1}$ remains an i.i.d. uniform family in each $\mathcal{Q}_k$. Let us denote by $\gamma_k(i,j)$ the distance between η_i and η_j in $\mathcal{Q}_k$ for $k \geq 1$. The main tool towards Theorem 5.8 is the following proposition:

Proposition 5.10. *For each $i, j \geq 1$, $\gamma_k(i,j) \to \gamma_\infty(i,j) < \infty$ almost surely as $k \to \infty$.*

Let us first explain why this entails Theorem 5.8.

Proof of Theorem 5.8. Observe that by Proposition 5.7, for each $k \geq 1$, $\mathcal{Q}_k$ is distributed as the Brownian CRT $\mathcal{H}$. Thus, $(\gamma_k(i,j))_{i,j\geq 1}$ and $(d_{\mathcal{H}}(\eta_i,\eta_j))_{i,j\geq 1}$ have the same distribution for each $k \geq 1$. It follows from Proposition 5.10 that the limit matrix $(\gamma_\infty(i,j))_{i,j\geq 1}$ is also distributed as $(d_{\mathcal{H}}(\eta_i,\eta_j))_{i,j\geq 1}$. In other words, $(\gamma_\infty(i,j))_{i,j\geq 1}$ has the distribution of the distance matrix of the Brownian CRT. In particular, for each $n \geq 1$, $(\gamma_\infty(i,j))_{1\leq i,j\leq n}$ defines an n-leaf real tree $\mathcal{R}_n$ and the family $(\mathcal{R}_n)_{n\geq 1}$ is *consistent* and *leaf-tight* (see [10]), which means that $(\mathcal{R}_n)_{n\geq 1}$ admits a representation as a continuum random tree.

Observe that all these are still true a.s. conditionally on $\mathcal{H}$. More precisely, by Proposition 5.10, for $\mathbf{P}$-almost every $\mathcal{H}$, as $k \to \infty$,

$$\big(\gamma_k(i,j), i,j \leq n\big) \xrightarrow{a.s.} \big(\gamma_\infty(i,j), i,j \leq n\big). \tag{5.14}$$

for each $n \geq 1$, conditionally on $\mathcal{H}$. For those $\mathcal{H}$ for which (5.14) holds, the family $(\mathcal{R}_n)_{n\geq 1}$ is a.s. consistent and leaf-tight, conditionally on $\mathcal{H}$. Let $\mathcal{R}_\infty(\mathcal{H})$ be the CRT representation of this family. Then by definition, (5.14) entails the convergence of $\mathrm{shuff}(\mathcal{H}, U_1, \cdots, U_k)$ to $\mathcal{R}_\infty(\mathcal{H})$ in Gromov–Prokhorov topology. $\square$

The remainder of the section is devoted to proving Proposition 5.10. Since $(\eta_i)_{i\geq 1}$ is an i.i.d. sequence, it suffices to consider the case $i = 1, j = 2$.

5.3.2 A series representation for $\gamma_k(1,2)$

The idea behind the formal definition of Proposition 5.6 is to leverage Proposition 5.5 as follows: if $\mathcal{H}$ were $\mathrm{cut}(\mathcal{T}, V_1)$, for some $\mathcal{T}$ and V_1, and the distinguished path were the one between the root and $U_1 \in \mathcal{H}$, then the image of a path between two points ξ_1 and ξ_2 in $\mathcal{T}$ would now go through a number of subtrees of $\mathcal{H} \setminus \mathrm{Span}(\mathcal{H}, U_1)$, and in every such tree it would go between two points which are uniform. We now go further and give such a representation for $\gamma_k(1,2)$ as a sum where we specify the distributions of the trees and points involved.

The masses of these trees are of prime importance, and we let

$$\mathcal{S}^\downarrow = \left\{(x_0, x_1, \cdots) : x_0 \geq x_1 \geq \cdots \geq 0; \sum_{i\geq 0} x_i \leq 1\right\}$$

be the space of mass partitions, equipped with the usual ℓ_1-norm $\|\cdot\|_1$. If $\mathbf{x} = (x_1, x_2, \cdots) \in \mathcal{S}^\downarrow$, then the *length* of $\mathbf{x}$ is defined to be the smallest index n such that $x_n = 0$, which may well be infinite. And we denote by $\mathcal{S}^\downarrow_f$ the subset of $\mathcal{S}^\downarrow$ which consists of the elements of finite length.

Recall the definition of $\gamma_1(1,2)$ in (5.7). The trees involved there are the components $F_{x_i(n)}$ rooted at $x_i(n)$, for $n \geq 0$ and $i = 1, 2$. Let ϖ denote the distribution of the rearrangement of

$$\{\nu(F_{x_1(n)}), 0 \leq n \leq \mathscr{I}(1,2)\} \cup \{\nu(F_{x_2(n)}), 0 \leq n \leq \mathscr{I}(2,1) - 1\}$$

in decreasing order. Then ϖ is a probability measure supported on $\mathcal{S}^\downarrow_f$.

Lemma 5.11 (Representation of $\gamma_k(1,2)$)**.** *For each $k \geq 1$, there exists a finite sub-collection of the masses of the components of $\mathcal{H} \setminus \mathrm{Span}(\mathcal{H}; U_1, \cdots, U_k)$ denoted by $\mathbf{m}_k = (m_{k,n})_{0 \leq n \leq N_k} \in \mathcal{S}_f^{\downarrow}$, and a sequence of positive real numbers $(\mathrm{R}_n^k, 0 \leq n \leq N_k)$ such that*

$$\gamma_k(1,2) = \sum_{n=0}^{N_k} \sqrt{m_{k,n}}\, \mathrm{R}_n^k, \tag{5.15}$$

where $(\mathrm{R}_n^k)_{n \geq 0}$ is a sequence of i.i.d. copies of a Rayleigh random variable that is independent of $\mathbf{m}_k$ and N_k. Moreover, $(\mathbf{m}_k)_{k \geq 1}$ is a Markov chain with initial law ϖ and the following transitions:

- *with probability $1 - \|\mathbf{m}_k\|_1$, $\mathbf{m}_{k+1} = \mathbf{m}_k$, and*
- *for $0 \leq n \leq N_k$, with probability $m_{k,n}$, $\mathbf{m}_{k+1}$ is obtained by replacing in $\mathbf{m}_k$ the element $m_{k,n}$ by $m_{k,n} \cdot \widetilde{\mathbf{m}}$, where $\widetilde{\mathbf{m}}$ has distribution ϖ, and then resorting the sequence in decreasing order.*

Proof. Let us first consider the case $k = 1$. Let $\{\mathrm{T}_{1,n}, 0 \leq n \leq N_1\}$ be the collection $\{F_{x_1(m)}, 0 \leq m \leq \mathscr{I}(1,2)\} \cup \{F_{x_2(m)}, 0 \leq m \leq \mathscr{I}(2,1)\}$ sorted in such a way that $\mathbf{m}_1 := (\nu(\mathrm{T}_{1,n}), 0 \leq n \leq N_1)$ is nonincreasing. We denote by $\mathbf{F}_1 = (\mathrm{T}_{1,n}, n \geq 0)$ and $\mathbf{m}_1 = (m_{1,n}, n \geq 0)$. Then the distribution of $\mathbf{m}_1$ is ϖ. For each $0 \leq n \leq N_1$, define (see Figure 5.4)

$$D_n^1 = \begin{cases} d_{\mathcal{H}}(x_1(m), a_1(m)), & \text{if } \mathrm{T}_{1,n} = F_{x_1(m)} \text{ for } m = 0, 1, \cdots, \mathscr{I}(1,2) - 1; \\ d_{\mathcal{H}}(x_2(m), a_2(m)), & \text{if } \mathrm{T}_{1,n} = F_{x_2(m)} \text{ for } m = 0, 1, \cdots, \mathscr{I}(2,1) - 1; \\ d_{\mathcal{H}}(a_1(\mathscr{I}(1,2)), a_2(\mathscr{I}(2,1))), & \text{otherwise.} \end{cases}$$

Then, $m_{1,n} > 0$ for $0 \leq n \leq N_1$, and we set $\mathrm{R}_n^1 := D_n^1 m_{1,n}^{-1/2}$. By (5.7), this definition immediately yields

$$\gamma_1(1,2) = \sum_{n=0}^{N_1} \sqrt{m_{1,n}}\, \mathrm{R}_n^1. \tag{5.16}$$

This is (5.15) for $k = 1$. For the distribution of $(\mathrm{R}_n^1)_{n \geq 1}$, we need the following fact, whose proof is given in Section 5.5.

Lemma 5.12 (Scaling property)**.** *For $0 \leq n \leq N_1$, let T_n^* be the rescaled metric space $m_{1,n}^{-1/2}\mathrm{T}_{1,n}$, equipped with the (probability rescaled) restriction of ν to $\mathrm{T}_{1,n}$. Then for each $j \geq 1$, on the event $\{N_1 = j\}$, $(\mathrm{T}_n^*)_{0 \leq n \leq N_1}$ is a sequence of j independent copies of a Brownian CRT.*

Let us recall that by definition each $a_i(m)$ is distributed as ν restricted to $F_{x_i(m)}$. We also recall that if η, η' are two independent points of $\mathcal{H}$ sampled according to ν, then $d_{\mathcal{H}}(\eta, \eta')$ (resp. the distance of η from the root) is Rayleigh distributed. Then it follows from Lemma 5.12 that $(\mathrm{R}_n^1)_{0 \leq n \leq N_1}$ is a sequence of i.i.d. Rayleigh random variables. This proves the statement for $k = 1$, which is our base case.

Now we proceed to prove the induction step, and assume that we almost surely have the desired representation for all natural numbers up to $k \geq 1$. In particular, there exists a sequence $\mathbf{F}_k = (\mathrm{T}_{k,n})_{0 \leq n \leq N_k}$ which is a finite sub-collection of the connected components of $\mathcal{H} \setminus \mathrm{Span}(\mathcal{H}; U_1, \cdots, U_k)$ such that $\mathbf{m}_k := (\nu(\mathrm{T}_{k,n}))_{0 \leq n \leq N_k}$ is non-increasing. Moreover, we suppose that (5.15) holds for k, where for each $0 \leq n \leq N_k$, $\sqrt{m_{k,n}}\mathrm{R}_n^k$ is either the distance between two independent ν-points of $\mathrm{T}_{k,n}$ or the distance between a ν-point and the root of $\mathrm{T}_{k,n}$.

Recall that $\mathcal{Q}_{k+1}$ is defined to be $\mathrm{shuff}(\tilde{\mathcal{H}}_{k+1}, U_1, \cdots, U_k)$, where $\tilde{\mathcal{H}}_{k+1}$ is obtained from $\mathcal{H}$ by replacing the connected component $\tilde{T}_{k+1}$ of $\mathcal{H} \setminus \mathrm{Span}(\mathcal{H}; U_1, \ldots, U_k)$ which contains U_{k+1}, by $\tilde{S} := \mathrm{shuff}(\tilde{T}_{k+1}, U_{k+1})$. The real tree $\tilde{\mathcal{H}}_{k+1}$ is a Brownian CRT with mass measure $\tilde{\nu}_{k+1}$, and by the induction hypothesis, there exists a sequence $\hat{\mathbf{F}}_k = (\hat{\mathrm{T}}_{k,n}, 0 \leq n \leq \hat{N}_k)$ which consists in a finite sub-collection of the components of $\tilde{\mathcal{H}}_{k+1} \setminus \mathrm{Span}(\tilde{\mathcal{H}}_{k+1}; U_1, \cdots, U_k)$ rearranged in decreasing order of their masses, such that

$$\gamma_{k+1}(1,2) = \sum_{n=0}^{\hat{N}_k} \sqrt{\hat{m}_{k,n}}\, \hat{\mathrm{R}}_n^k, \tag{5.17}$$

where for each $0 \le n \le \hat{N}_k$, $\hat{m}_{k,n} = \mu(\hat{\mathrm{T}}_{k,n})$ and $\sqrt{\hat{m}_{k,n}}\hat{\mathrm{R}}_n^k$ is either the distance between two uniform independent points of $\hat{\mathrm{T}}_{k,n}$ or the distance between a uniform point and the root of $\hat{\mathrm{T}}_{k,n}$. However, we are after a representation in terms of the connected components of $\mathcal{H} \setminus \mathrm{Span}(\mathcal{H}; U_1, \dots, U_{k+1})$.

Note that $\tilde{S}$ is the only component of $\tilde{\mathcal{H}}_{k+1} \setminus \mathrm{Span}(\tilde{\mathcal{H}}_{k+1}; U_1, \cdots, U_k)$ that is not a component of $\mathcal{H}\setminus\mathrm{Span}(\mathcal{H}; U_1, \dots, U_{k+1})$. So if $\tilde{S}$ does not appear in $\hat{\mathbf{F}}_k$, then by construction we have $\hat{\mathbf{F}}_k = \mathbf{F}_k$ where $\mathbf{F}_k = (\mathrm{T}_{k,n})_{0\le n\le N_k}$ is a finite sub-collection of the connected components of $\mathcal{H}\setminus\mathrm{Span}(\mathcal{H}; U_1, \cdots, U_k)$ such that $\mathbf{m}_k := (\nu(\mathrm{T}_{k,n}))_{0\le n\le N_k}$ is non-increasing for which (5.15) holds, with the additional distributional properties we are after. Furthermore, in that case, we have $\gamma_{k+1}(1,2) = \gamma_k(1,2)$. It thus suffices to take $\mathbf{F}_{k+1} = \mathbf{F}_k$ and $\mathrm{R}_n^{k+1} = \mathrm{R}_n^k$ for each n. This case occurs precisely if U_{k+1} does not fall in any of the subtrees of $\mathbf{F}_k$, which happens with probability $1 - \|\mathbf{m}_k\|_1$, since U_{k+1} is ν-distributed.

If, on the other hand, $\tilde{S} = \hat{\mathrm{T}}_{k,n_0}$ for some $0 \le n_0 \le \hat{N}_k$, then the representation in (5.17) needs to be modified. Still, since $\hat{m}_{k,n_0} = \tilde{\nu}_{k+1}(\tilde{S}) = \nu(\tilde{T}_{k+1}) = m_{k,n_0}$, the masses are correct and $\hat{\mathbf{m}}_k = \mathbf{m}_k$. So, in particular, this occurs with probability m_{k,n_0} Note also that, by definition of $\tilde{\mathcal{H}}_{k+1}$,

$$\gamma_{k+1}(1,2) - \gamma_k(1,2) = \hat{\mathrm{R}}_{n_0}^k - \mathrm{R}_{n_0}^k, \tag{5.18}$$

where here, $(m_{k,n_0})^{1/2}\hat{\mathrm{R}}_{n_0}^k$ is the distance in $\tilde{S}$ between either two independent $\tilde{\nu}_{k+1}$-points or between a $\tilde{\nu}_{k+1}$-point and the root. Recall that $\tilde{S} = \mathrm{shuff}(\tilde{T}_{k+1}, U_{k+1})$ is rooted at a $\tilde{\nu}_{k+1}$-point. Note also that by the scaling property, $(m_{k,n_0})^{-1/2}\tilde{T}_{k+1}$ is a Brownian CRT (and is thus distributed as $\mathcal{H}$). Therefore, we may use the induction hypothesis with $k = 1$ to obtain that there exists a sequence $\check{\mathbf{F}} = (\check{\mathrm{T}}_n, 0 \le n \le \check{N})$ consisting in a sub-collection of the connected components of $\tilde{T}_{k+1} \setminus \mathrm{Span}(\tilde{T}_{k+1}; U_{k+1})$ rearranged in the decreasing order of their masses such that

$$\hat{\mathrm{R}}_{n_0}^k = \sum_{n=0}^{\check{N}} \sqrt{\nu(\check{\mathrm{T}}_n)}\, \check{\mathrm{R}}_n, \tag{5.19}$$

where for each $0 \le n \le \check{N}$, $\nu(\check{T}_n)^{1/2}\check{\mathrm{R}}_n$ is either the distance in $\check{\mathrm{T}}_n$ between either two independent ν-points or between a ν-point and the root. So in particular, $(\check{\mathrm{R}}_n)_{n\ge0}$ forms a sequence of i.i.d. copies of a Rayleigh random variable. Furthermore, $(\nu(\check{\mathrm{T}}_n)/m_{k,n_0}, 0 \le n \le \check{N})$ is an independent copy of $\mathbf{m}_1$. Then we set $\mathbf{F}_{k+1} = (\mathrm{T}_{k+1,n}, 0 \le n \le N_{k+1})$ to be the rearrangement of the collection

$$\{\hat{\mathrm{T}}_{k,n} : 0 \le n \le \hat{N}_k, n \ne n_0\} \cup \{\check{\mathrm{T}}_n : 0 \le n \le \check{N}\}$$

such that $\mathbf{m}_{k+1} := (\nu(\mathrm{T}_{k+1,n}))_{0\le n\le N_{k+1}}$ is non-increasing. Note that $\mathbf{F}_{k+1}$ is a finite sub-collection of components of $\mathcal{H} \setminus \mathrm{Span}(\mathcal{H}; U_1, \cdots, U_{k+1})$. Finally, inserting (5.19) into (5.17) and then comparing with (5.18), we obtain (5.15) for $k + 1$, which completes the proof. □

Proving Proposition 5.10 now reduces to showing that the series representation in (5.15) converges almost surely. First observe that conditionally on $\mathbf{m}_k$, $\gamma_k(1,2)$ is a sum of independent random variables which have Gaussian tails (see later for details). It then easily follows from classical results on concentration of measure that $\gamma_k(1,2)$ is concentrated about the conditional mean $\mathbb{E}[\gamma_k(1,2) \mid \mathbf{m}_k]$. Furthermore the width of the concentration window is controlled by the variance, which is here $O(\|\mathbf{m}_k\|_1)$. The following lemmas control the distance between $\gamma_k(1,2)$ and $\mathbb{E}[\gamma_k(1,2) \mid \mathbf{m}_k]$.

Lemma 5.13. *There exists some $\alpha > 0$ such that*

$$\lim_{k\to\infty} k^\alpha \|\mathbf{m}_k\|_1 = 0, \qquad \textit{almost surely.}$$

Lemma 5.14. *We have $\gamma_k(1,2) - \mathbb{E}[\gamma_k(1,2) \,|\, \mathbf{m}_k] \to 0$ almost surely as $k \to \infty$.*

The proofs of these lemmas rely on standard facts about fragmentations chains and concentration inequalities and are presented in Sections 5.3.3 and 5.3.4. From there, the last step consists in proving that $\mathbb{E}[\gamma_k(1,2) \,|\, \mathbf{m}_k]$ also converges almost surely.

Lemma 5.15. *A.s., $\mathbb{E}[\gamma_k(1,2) \,|\, \mathbf{m}_k]$ converges to some random variable $\gamma_\infty(1,2) < \infty$ as $k \to \infty$.*

Our approach to Lemma 5.15 relies on a coupling between the cutting and shuffling procedure and is given in Section 5.3.5.

5.3.3 Proof of Lemma 5.13: polynomial decay of the self-similar fragmentation chain

The dynamics of $(\mathbf{m}_k)_{k\geq 0}$ are quite similar to that of a self-similar fragmentation chain, and the proof of the lemma relies on classical results on the asymptotic behaviour of fragmentation processes. If we were to count only the number of *actual* jumps of the process (or equivalently the number of i's such that U_i does affect the collection of masses) then one would exactly have the state of a fragmentation chain taken at the jump times. The fact that even the i's that do not affect the chain are counted only induces a time-change that is easily controlled.

Recall that ϖ denotes the law of $\mathbf{m}_1$. Let $X(t) = (X_i(t))_{i\geq 1}$ be a self-similar fragmentation chain with index of self-similarity 1 and dislocation measure ϖ starting from the initial state $X(0) = (1, 0, 0, \cdots)$, as introduced in [26, Chapter 1]. Then, $X(t)$ jumps at rate $\|X(t)\|_1 = \sum_{i\geq 1} X_i(t)$. This chain is non-conservative since $\mathbf{P}(0 < \|\mathbf{m}_1\|_1 < 1) = 1$ for $\mathbf{m}_1$ is an a.s. finite and non-empty collection of the masses of the components of $\mathcal{H} \setminus \mathrm{Span}(\mathcal{H}; U_1)$. To compensate for the loss of mass (the i's that do not modify $(\mathbf{m}_k)_{k\geq 0}$), consider another Poisson point process Γ on $[0,\infty)$ with rate $1 - \|X(t)\|_1$ at time $t \geq 0$, and let $\theta(t)$ denote the number of jumps before time t in X and Γ combined. Then $\theta(t)$ is the number of jumps before time t of a Poisson process with rate one, and if we set $\theta^{-1}(k) := \inf\{t \geq 0 : \theta(t) \geq k\}$ be the time of the k-th point, then we have

$$(X(\theta^{-1}(k)))_{k\geq 1} \stackrel{d}{=} (\mathbf{m}_k)_{k\geq 1}. \tag{5.20}$$

From now on, we work on a space on which these are coupled to be equal with probability one.

Let $p^* \in (0,1)$ be the critical exponent such that $\mathbb{E}[\sum_{i\geq 1} m_{1,i}^{p^*}] = 1$. Then

$$\mathcal{M}(t) := \sum_{i\geq 1} X_i^{p^*}(t)$$

is a positive martingale which is uniformly integrable. Denote by $\mathcal{M}(\infty)$ its a.s. limit. By Theorem 1 of [29], for every $k \geq 1$, there exists a constant $C_k \in (0,\infty)$ such that

$$\sup_{t\geq 0} t^k \cdot \mathbb{E}\left[\sum_{i\geq 1} X_i^{p^*+k}(t)\right] \leq C_k,$$

from which it follows immediately that $\sup_{t\geq 0} t^k \cdot \mathbb{E}X_1^{p^*+k}(t) \leq C_k$. Now, for any $\delta \in (0,1)$ and $\epsilon > 0$, by Markov's inequality, we obtain at time $n \geq 0$

$$\mathbf{P}\big(X_1(n) \geq \epsilon n^{-\delta}\big) \leq \frac{n^k \cdot \mathbb{E}X_1^{p^*+k}(n)}{\epsilon^{p^*+k} n^{(1-\delta)k-\delta p^*}} \leq \frac{C_k}{\epsilon^{p^*+k} n^{(1-\delta)k-\delta p^*}}.$$

By choosing k large enough that $k > (1+\delta p^*)/(1-\delta)$, this implies that $\sum_{n\geq 1} \mathbf{P}(X_1(n) \geq \epsilon n^{-\delta}) < \infty$ and $\limsup_{n\to\infty} n^\delta X_1(n) \leq \epsilon$ almost surely, by the Borel–Cantelli lemma. Letting $\epsilon \to 0$ along a sequence, we then obtain that $n^\delta X_1(n) \to 0$ a.s., and $t^\delta X_1(t) \to 0$ a.s. as well by monotonicity. Now notice that for any $t \geq 0$,

$$\sum_{i\geq 1} X_i(t) = \sum_{i\geq 1} X_i^{1-p^*}(t) \cdot X_i^{p^*}(t) \leq X_1^{1-p^*}(t) \cdot \mathcal{M}(t).$$

Then, for any $\delta \in (0,1)$,

$$\limsup_{t\to\infty} t^{\delta(1-p^*)} \sum_{i\geq 1} X_i(t) \leq \mathcal{M}(\infty) \cdot \limsup_{t\to\infty} (t^\delta X_1(t))^{1-p^*} = 0 \quad \text{almost surely,} \tag{5.21}$$

since $p^* \in (0,1)$ as the chain is non-conservative.

Finally, we go back to $(\mathbf{m}_k)_{k\geq 1}$ using (5.20). By the strong law of large numbers, we have $\theta(t)/t \to 1$ almost surely, and we easily deduce that $k/\theta^{-1}(k) \to 1$ almost surely as $k \to \infty$. Using this fact together with (5.20) and (5.21), we obtain that for any $\delta \in (0,1)$,

$$\limsup_{k\to\infty} k^{\delta(1-p^*)} \|\mathbf{m}_k\|_1 = \limsup_{k\to\infty} k^{\delta(1-p^*)} \sum_{i\geq 1} X_i(\theta^{-1}(k)) = 0 \quad \text{almost surely,}$$

which completes the proof of Lemma 5.13.

5.3.4 Proof of Lemma 5.14: concentration of the Rayleigh variable

Let R be a random variable of Rayleigh distribution, with density $xe^{-x^2/2}$ on $\mathbb{R}_+$. Then, one easily verifies that $\mathrm{R} - \mathbb{E}[\mathrm{R}]$ is *sub-Gaussian* in the sense that there exists a constant v such that for every $\lambda \in \mathbb{R}$, one has

$$\log \mathbb{E}[e^{\lambda(\mathrm{R}-\mathbb{E}[\mathrm{R}])}] \leq \frac{\lambda^2 v}{2}.$$

(See [36, Theorem 2.1, p. 25].) We may thus apply concentration results for sub-Gaussian random variables such as those presented in Section 2.3 of [36].

For each $k \geq 1$, we have

$$\sigma_k := \gamma_k(1,2) - \widetilde{\gamma}_k = \sum_{i=1}^{N_k} \sqrt{m_{k,n}}\left(\mathrm{R}_n^k - \mathbb{E}[\mathrm{R}_n^k]\right),$$

by (5.15), where according to Lemma 5.11, $(\mathrm{R}_n^k, 1 \leq n \leq N_k)$ is a sequence of i.i.d. copies of a Rayleigh random variable. Therefore

$$\mathbf{P}\left(|\sigma_k| \geq \epsilon \,|\, \mathbf{m}_k\right) \leq 2\exp\left(-\frac{\epsilon^2}{2v\|\mathbf{m}_k\|_1}\right). \tag{5.22}$$

If $(A_k, B_k, C_k)_{k\geq 1}$ are sequence of events satisfying that $A_k \subset B_k \cup C_k$ for each $k \geq 1$, then it is elementary that $\mathbf{P}(\limsup_k A_k) \leq \mathbf{P}(\limsup_k B_k) + \mathbf{P}(\limsup_k C_k)$. Here, we take

$$A_k = \{|\sigma_k| \geq \epsilon\}, \quad B_k = A_k \cap \{\|\mathbf{m}_k\|_1 \leq k^{-\alpha}\}, \quad C_k = \{\|\mathbf{m}_k\|_1 > k^{-\alpha}\}$$

with the same α as in Lemma 5.13. Then $\mathbf{P}(\limsup_k C_k) = 0$ by Lemma 5.13. On the other hand, we deduce from (5.22) that

$$\sum_{k\geq 1} \mathbf{P}(B_k) = \sum_{k\geq 1} \mathbb{E}\left[\mathbf{P}\left(|\sigma_k| \geq \epsilon \,|\, \mathbf{m}_k\right) \cdot \mathbf{1}_{\{\|\mathbf{m}_k\|_1 \leq k^{-\alpha}\}}\right] \leq \sum_{k\geq 1} 2e^{-\epsilon^2 k^\alpha/(2v)} < \infty,$$

which entails that $\mathbf{P}(\limsup_k B_k) = 0$ by the Borel–Cantelli lemma. Hence, $\mathbf{P}\left(\limsup_k A_k\right) = 0$, which means $\limsup_k |\sigma_k| < \epsilon$ almost surely. Since $\epsilon > 0$ was arbitrary, the proof of Lemma 5.14 is now complete.

5.3.5 Proof of Lemma 5.15: a coupling via cut trees

Let us recall the notations before Proposition 5.5. There are two μ-points ξ_1, ξ_2 in the Brownian CRT $\mathcal{T}$, and $p :=]\!]\xi_1, \xi_2[\![$ is the path in $\mathcal{T}$ between these two points. We denote by $D := d_{\mathcal{T}}(\xi_1, \xi_2)$ the length of this path. Let $\mathcal{G}_k = \mathrm{cut}(\mathcal{T}, V_1, \cdots, V_k)$ be the k-cut tree. Recall that, up to the finitely many cut points that are lost, p_k is the image of p by the canonical embedding ϕ_k from $\cup_{t\in\mathcal{C}^k}\Delta_t^k$ into $\mathcal{G}_k$. We have the following representation of the distance D, which is an analog of Lemma 5.11.

Lemma 5.16. *For each $k \geq 1$, there exists some $\mathbf{m}'_k = (m'_{k,n})_{0\leq n\leq N'_k} \in \mathcal{S}_f^{\downarrow}$, which is a sub-collection of the masses of $\{\Delta_t^k, t \in \mathcal{C}^k\}$ such that*

$$D = \sum_{n=0}^{N'_k} \sqrt{m'_{k,n}}\, \mathrm{B}_n^k, \tag{5.23}$$

where $(\mathrm{B}_n^k)_{n\geq 0}$ is an i.i.d. sequence of Rayleigh random variables, independent of $\mathbf{m}'_k$ and N'_k. Moreover, $(\mathbf{m}'_k)_{k\geq 1}$ has the same distribution as $(\mathbf{m}_k)_{k\geq 1}$.

Proof. For each $k \geq 1$, the injection ϕ_k is an isometry on each $\mathrm{Sk}(\Delta_t^k)$. Thus, we have $D = \ell(p_k)$ for each $k \geq 1$. Let us show that $\ell(p_k)$ can be written as the right-hand side in (5.23).

We proceed by induction on $k \geq 1$. The base case $k = 1$ is a consequence of Proposition 5.5. Let $\mathbf{F}'_1 = (\mathrm{T}'_{1,n}, 0 \leq n \leq N'_1)$ be the vector consisting of the elements of the collection

$$\{\Delta^1_{t_{1,m}}, 0 \leq m \leq M_1\} \cup \{\Delta^1_{t_{2,m}}, 0 \leq m \leq M_2 - 1\},$$

arranged in the decreasing order of their masses $m'_{1,n} := \mu(\mathrm{T}'_{1,n})$. By comparing Propositions 5.5 and 5.6, we see that $\mathbf{F}'_1$ has the same distribution as $\mathbf{F}_1$, since $\mathcal{G}_1$ is a Brownian CRT (Proposition 5.1) and given $\mathcal{G}_1$ the sequences $(a'_i(m))_{m\geq 0}$, $i \geq 1$, are sampled in the same way as $(a_i(m))_{m\geq 0}$ given $\mathcal{H}$. As a consequence, $\mathbf{m}'_1 := (m'_{1,n}, 0 \leq n \leq N'_1)$ has the same distribution as $\mathbf{m}_1$. Moreover, by Lemma 5.12, each $(m'_{1,n})^{-1/2}\mathrm{T}'_{1,n}$ is an independent copy of a Brownian CRT. Define

$$D_n^1 := \begin{cases} d_{\mathcal{T}}(x'_1(m), a'_1(m)) & \text{if } \mathrm{T}'_{1,n} = \Delta^1_{t_{1,m}} \text{ for } m = 0, 1, \cdots, M_1 - 1; \\ d_{\mathcal{T}}(x'_2(m), a'_2(m)), & \text{if } \mathrm{T}'_{1,n} = \Delta^1_{t_{2,m}} \text{ for } m = 0, 1, \cdots, M_2 - 1; \\ d_{\mathcal{T}}(a'_1(M_1), a'_2(M_2)), & \text{otherwise,} \end{cases}$$

and set $\mathrm{B}_n^1 := D_n^1/(m'_{1,n})^{1/2}$. Then B_n^1, $0 \leq n \leq N'_1$, are independent from each other. Also, it follows from (5.5) that

$$\ell(p_1) = \sum_{n=0}^{N'_1} (m'_{1,n})^{1/2}\, \mathrm{B}_n^1.$$

Suppose now that for all natural numbers up to some $k \geq 1$ there exist some $\mathbf{F}'_k = (\mathrm{T}'_{k,n}, 0 \leq k \leq N'_k)$, the elements of which form a sub-collection of $(\Delta_t^k, t \in \mathcal{C}^k)$, such that $\mathbf{m}'_k = (m'_{k,n})_{0\leq n\leq N'_k} := (\mu(\mathrm{T}'_{k,n}))_{0\leq n\leq N'_k}$ is non-increasing and that

$$\ell(p_k) = \sum_{n=0}^{N'_k} \sqrt{m'_{k,n}}\, \mathrm{B}_n^k, \tag{5.24}$$

where $(m_{k,n})^{1/2}\mathrm{B}_n^k$ is the distance between two points, say u_n^k and v_n^k, in $\mathrm{T}'_{k,n}$, and u_n^k is a μ-point in $\mathrm{T}'_{k,n}$ while v_n^k is either the root of $\mathrm{T}'_{k,n}$ or another μ-point independent of u_n^k. Recall that by Proposition 5.3, $\mathcal{G}_{k+1}$ can be obtained from $\mathcal{G}_k$ by replacing $\Delta^k_{\tau_{k+1}}$ with $\mathrm{cut}(\Delta^k_{\tau_{k+1}}, V_{k+1})$.

Suppose first that $\Delta^k_{\tau_{k+1}}$ does not appear in $\mathbf{F}'_k$, which happens with probability $1-\|\mathbf{m}'_k\|_1$. Then, all the components of $\mathbf{F}'_k$ are actually elements of $\{\Delta^{k+1}_t, t\in\mathcal{C}^{k+1}\}$ (see Figure 5.3) and it suffices to take $\mathbf{F}'_{k+1}=\mathbf{F}'_k$ and $\mathrm{B}^{k+1}_n=\mathrm{B}^k_n$ for each n. So in this case, the representation in (5.23) for $k+1$ follows trivially from (5.24).

Suppose now that $\Delta^k_{\tau_{k+1}}=\mathrm{T}'_{k,n_0}$ for some $0\le n_0\le N'_k$, which occurs with probability m'_{k,n_0}. In this case, we have

$$\ell(p_{k+1})-\ell(p_k)=\ell(\tilde p)-\sqrt{m'_{k,n_0}}\,\mathrm{B}^k_{n_0}, \tag{5.25}$$

where $\tilde p := p_{k+1}\cap \mathrm{cut}(\Delta^k_{\tau_{k+1}},V_{k+1})$ is the image of $]\!]u^k_{n_0},v^k_{n_0}[\![\subset\Delta^k_{\tau_{k+1}}$ in $\mathrm{cut}(\Delta^k_{\tau_{k+1}},V_{k+1})$. Observe that the root behaves as a uniform point in our cutting procedure, and that the rescaled tree $(m'_{k,n_0})^{-1/2}\Delta^k_{\tau_{k+1}}$ is a standard Brownian CRT. Thus, the induction hypothesis for $k=1$ applies and with probability one there exists a sequence $(\tilde{\mathrm{T}}'_n, 0\le n\le\tilde N')$, which is a sub-collection of the Δ^{k+1}_t, $t\in\mathcal{C}^{k+1}$, which are subsets of $\mathrm{cut}(\Delta^k_{\tau_{k+1}},V_{k+1})$ (see Figure 5.3), rearranged in the decreasing order of their masses such that

$$\ell(\tilde p)=\sum_{n=0}^{\tilde N'}\sqrt{\mu(\tilde{\mathrm{T}}'_n)}\,\tilde{\mathrm{R}}'_n, \tag{5.26}$$

where $\mu(\tilde{\mathrm{T}}'_n)^{1/2}\tilde{\mathrm{R}}'_n$ is either the distance between two uniform independent points of $\tilde{\mathrm{T}}'_n$ or the distance between a uniform point and the root of $\tilde{\mathrm{T}}'_n$, so that $(\tilde{\mathrm{R}}'_n)$ is an i.i.d. family of Rayleigh distributed random variables. Furthermore, $(\mu(\tilde{\mathrm{T}}'_n)/m'_{k,n_0}, 0\le n\le\tilde N')$ is an independent copy of $\mathbf{m}_1$. Then we set $\mathbf{F}'_{k+1}=(\mathrm{T}'_{k+1,n},0\le n\le N'_{k+1})$ to be the rearrangement of the collection

$$\{\mathrm{T}'_{k,n}:0\le n\le N'_k, n\ne n_0\}\cup\{\tilde{\mathrm{T}}'_n:0\le n\le\tilde N'\}$$

such that $\mathbf{m}'_{k+1}:=(\mu(\mathrm{T}'_{k+1,n}),0\le n\le N'_{k+1})$ is non-increasing. Inserting (5.25) and (5.26) into (5.24) yields the representation in (5.23) for $k+1$, which completes the proof of the induction step. $\square$

As $(\mathbf{m}'_k)_{k\ge1}$ has the same distribution as $(\mathbf{m}_k)_{k\ge1}$, Lemma 5.13 also holds for $(\mathbf{m}'_k)_{k\ge1}$. Furthermore the concentration arguments already used in the course of the proof of Lemma 5.14 imply that, a.s.,

$$D-\mathbb{E}[D\,|\,\mathbf{m}'_k]=D-\sqrt{\pi/2}\sum_{n=0}^{N'_k}(m'_{k,n})^{1/2}=\sum_{n=0}^{N'_k}(m'_{k,n})^{1/2}\cdot\left(\mathrm{B}^k_n-\mathbb{E}[\mathrm{B}^k_n]\right)\to0.$$

Since D does not vary with k, this implies that $(\mathbb{E}[D\,|\,\mathbf{m}'_k])_{k\ge1}$ converges almost surely. Since the sequence $(\mathbb{E}[\gamma_k(1,2)\,|\,\mathbf{m}_k])_{k\ge1}$ has the same distribution, it also converges almost surely and the proof of Lemma 5.15 is complete.

5.4 Direct construction of the complete reversal $\mathrm{shuff}(\mathcal{H})$

In this section, we finally prove that the operation which we described in the introduction as the dual to the complete cutting procedure makes sense, and is indeed the desired dual. This reduces to make the link between the collection of random variables $(A_x, x\in\mathrm{Br}(\mathcal{T}))$ and the iterative reversal of paths to the random leaves $(U_i)_{i\ge1}$. We prove here that the sequence of random leaves can be constructed as a measurable function of the family $(A_x,x\in\mathrm{Br}(\mathcal{T}))$.

5.4.1 Construction of one consistent leaf

Recall that $\mathcal{H}$ is a Brownian CRT rooted at ρ and with mass measure ν. Let $\{A_x, x \in \mathrm{Br}(\mathcal{H})\}$ be a family of independent random variables such that for every x, the point $A_x \in \mathrm{Sub}(\mathcal{H}, x)$ is chosen according to the restriction of the mass measure ν to $\mathrm{Sub}(\mathcal{H}, x)$.

One of the main constraints for the family $(U_i)_{i\geq 1}$ is that the path to U_1 should be the first path to be reversed, and because of this, all the branch points x on the path between the root and U_1 should have choices A_x which are consistent with the reversing of the path to U_1 in the sense that for every branch point x on $[\![\rho, U_1]\!]$, one should have

$$\mathrm{Fr}(\mathcal{H}, x, A_x) = \mathrm{Fr}(\mathcal{H}, x, U_1).$$

Think of the discrete setting: if we have a rooted tree T and $(A_u, u \in T)$ a sequence of independent nodes such that A_u is distributed uniformly in the tree above u, then it is easy to construct a uniformly random node U_1 such that for every node u on the path between U_1 and the root, $[\![u, U_1]\!] \cap [\![u, A_u]\!] \setminus \{u\} \neq \varnothing$, or equivalently the points U_1 and A_u both lie in the same subtree of T rooted at one of the children of u. To do this, one simply needs to build the path from the root to U_1 by iteratively adding nodes as follows: start from the root; at some step where the current node is v, if $A_v \neq v$ then move to the first node in the direction to A_v, otherwise $A_v = v$ and set $U_1 = v$. The node U_1 is constructed so that the choices on the path $[\![\rho, U_1]\!]$ are consistent, and one easily verifies that U_1 is a uniformly random node. The idea is to adapt this technique to the continuous setting. To this aim, it suffices to verify that, when constructing a consistent path from the root, we make positive progress.

Let $v \in \mathcal{H} \setminus \{\rho\}$. We say that a branch point $x \in [\![\rho, v]\!] \cap \mathrm{Br}(\mathcal{H})$ is a *turning point* between the root and v if $A_x \in \mathrm{Fr}(\mathcal{H}, x, v)$. We denote by $\mathcal{N}_v$ the set of those x which are turning points between the root and v.

Lemma 5.17. *Let $\mathcal{H}$ be the Brownian CRT. For each $y \in \mathrm{Sk}(\mathcal{H})$, with probability one, $\mathcal{N}_y$ has at most finitely many elements.*

Proof. Let us write $\nu_y = \nu(\mathrm{Sub}(\mathcal{H}, y))$. As $\mathcal{H}$ is a Brownian CRT $\nu_y > 0$ almost surely for every $y \in \mathrm{Sk}(\mathcal{H})$. Let us write $\mathbf{P}_{\mathcal{H}}$ for the probability measure conditionally on $\mathcal{H}$. Since A_x is distributed according to the restriction of ν to $\mathrm{Sub}(\mathcal{H}, x)$, we have for each $x \in [\![\rho, y]\!] \cap \mathrm{Br}(\mathcal{H})$,

$$\mathbf{P}_{\mathcal{H}}(x \in \mathcal{N}_y) = \frac{\nu(\mathrm{Fr}(\mathcal{H}, x, y))}{\nu(\mathrm{Sub}(\mathcal{H}, x))} \leq \frac{\nu(\mathrm{Fr}(\mathcal{H}, x, y))}{\nu_y}$$

Then, we have

$$\begin{aligned}
\mathbb{E}_{\mathcal{H}}[\mathrm{Card}(\mathcal{N}_y)] &= \sum_{x \in [\![\rho, y]\!] \cap \mathrm{Br}(\mathcal{H})} \mathbf{P}_{\mathcal{H}}(x \in \mathcal{N}_y) \\
&\leq \sum_{x \in [\![\rho, y]\!] \cap \mathrm{Br}(\mathcal{H})} \frac{\nu(\mathrm{Fr}(\mathcal{H}, x, y))}{\nu_y} = \frac{1}{\nu_y} < \infty, \quad \text{almost surely,}
\end{aligned}$$

since the different fringe trees are disjoint. This shows that almost surely $\mathrm{Card}(\mathcal{N}_y) < \infty$. □

Now let us consider the following iterative process which constructs refining approximations Y_k, $k \geq 0$, of U. Start with $Y_0 = \rho$ and Z_0, which is a leaf of distribution ν. Supposing now that we have defined Y_k and Z_k, for some $k \geq 0$, if $\mathcal{N}_{Z_k} = \varnothing$, then we stop the process and set $U = Z_k$. Otherwise, there must exist some $y \in [\![\rho, Z_k]\!] \cap \mathrm{Sk}(\mathcal{H})$ such that $\mathcal{N}_y \neq \varnothing$. But by Lemma 5.17, $\mathcal{N}_y$ is finite a.s. so that there is an $x_0 \in \mathcal{N}_y$ which is closest to the root. Then we set $Y_{k+1} = x_0$ and $Z_{k+1} = A_{Y_{k+1}}$. Note that $Y_k \notin \mathcal{N}_{Z_k}$, so $Y_{k+1} \in \mathrm{Sub}(\mathcal{H}, Y_k) \setminus \{Y_k\}$.

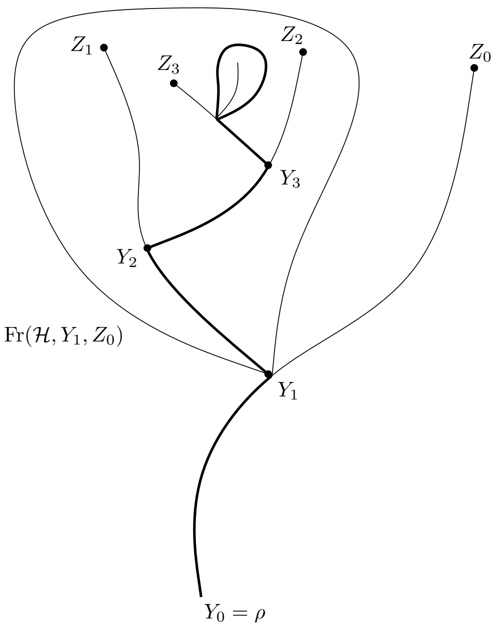

Figure 5.5 – *The iterative construction of the path $Y_1, Y_2, \ldots$ to the random leaf U.*

Lemma 5.18. *As $k \to \infty$, we have $Y_k, Z_k \to U$ almost surely. Furthermore, U is a ν-distributed leaf.*

Proof. Let us first show that Z_k converges to some point U. If the process has stopped at some finite time, this is obvious. Otherwise, note that $(\mathrm{Sub}(\mathcal{H}, Y_k))_{k\geq 1}$ is a decreasing sequence of sets and define U in $\cap_{k\geq 0} \mathrm{Sub}(\mathcal{H}, Y_k)$ to be the point which minimizes the distance to the root. (So $\cap_{k\geq 0} \mathrm{Sub}(\mathcal{H}, Y_k) = \mathrm{Sub}(\mathcal{H}, U)$ and $Y_k \to U$.) Now, we claim that U is a leaf, so that $Z_k \to U$ as $k \to \infty$. To see this, suppose for a contradiction that U is not a leaf. Then $U \in \mathrm{Sk}(\mathcal{H})$ and with probability one, $\nu^\star := \nu(\mathrm{Sub}(\mathcal{H}, U)) > 0$. Note that by construction, we have $Z_k \notin \mathrm{Fr}(\mathcal{H}, Y_{k+1}, Z_k) \supset \mathrm{Sub}(\mathcal{H}, Y_{k+2})$, thus $Z_k \notin \mathrm{Sub}(\mathcal{H}, U)$ for each $k \geq 1$. However, for every k,

$$\begin{aligned}\mathbf{P}_{\mathcal{H}}(Z_i \notin \mathrm{Sub}(\mathcal{H}, U), i = 0, \ldots, k) &\leq \prod_{i=0}^{k} \left(1 - \frac{\nu(\mathrm{Sub}(\mathcal{H}, U))}{\nu(\mathcal{H}, Y_k)}\right) \\ &\leq (1 - \nu^\star)^k \to 0,\end{aligned}$$

as $k \to \infty$. As a consequence, almost surely $\nu^\star = \nu(\mathrm{Sub}(\mathcal{H}, U)) = 0$ so that U is a.s. a leaf.

It now remains to prove that U is indeed ν-distributed. For this, it suffices to show that for every $x \in \mathcal{H}$, $\mathbf{P}_{\mathcal{H}}(U \in \mathrm{Sub}(\mathcal{H}, x)) = \nu(\mathrm{Sub}(\mathcal{H}, x))$. Note that since U is a leaf, we have $Z_k \to U$ as $k \to \infty$. We claim that for every $k \geq 0$, the leaf Z_k is ν distributed. Clearly, this would complete the proof. We proceed by induction on $k \geq 0$. For $k = 0$, Z_0 has distribution μ and the result is immediate. It will be useful to prove also the result for $k = 1$.

For a point $s \in [\![Y_0, Z_0]\!]$, we write $F_s = \mathrm{Fr}(\mathcal{H}, s, Z_0)$. For any $x \in [\![Y_0, Z_0]\!]$, since the branch points are countable, we have

$$\mathbf{P}_{\mathcal{H}}(Y_1 \in \mathrm{Sub}(\mathcal{H}, x) \,|\, Z_0, Y_0) = \prod_{s \in [\![Y_0, x[\![} \frac{\nu(\mathrm{Sub}(\mathcal{H}, s) \setminus F_s)}{\nu(\mathrm{Sub}(\mathcal{H}, s))} = \prod_{s \in [\![Y_0, x[\![} \left(1 - \frac{\nu(F_s)}{\nu(\mathrm{Sub}(\mathcal{H}, s))}\right), \quad (5.27)$$

since the choices of all the points are independent. There are only finitely $s \in [\![Y_0, x[\![$ for which $\nu(F_s) > \nu(\mathrm{Sub}(\mathcal{H}, x))/2$. For the others, $\nu(F_s)/\nu(\mathrm{Sub}(\mathcal{H}, s)) \leq 1/2$ and

$$1 - \frac{\nu(F_s)}{\nu(\mathrm{Sub}(\mathcal{H}, s))} \geq \exp\left(-2\frac{\nu(F_s)}{\mathrm{Sub}(\mathcal{H}, x))}\right).$$

It follows that the infinite product in (5.27) is absolutely convergent since $\sum_{s\in[\![Y_0,x]\!]}\nu(F_s)\le 1$. Therefore, we have $\mathbf{P}_{\mathcal{H}}(Y_1\in\mathrm{Sub}(\mathcal{H},x)\,|\,Z_0,Y_0)=\nu(\mathrm{Sub}(\mathcal{H},x))$. It follows that $\mathbf{P}_{\mathcal{H}}(Y_1=s\,|\,Z_0,Y_0)=\nu(F_s)$, for $s\in[\![Y_0,Z_0]\!]$. Note that from our definition of F_s, this is indeed a probability distribution since $\sum_{s\in[\![Y_0,Z_0]\!]}\nu(F_s)=1$. Now, for any $z\in\mathcal{H}$, we have

$$\mathbf{P}_{\mathcal{H}}(Z_1\in\mathrm{Sub}(\mathcal{H},z)\,|\,Y_1=x,Z_0)=\frac{\nu(\mathrm{Sub}(\mathcal{H},z))}{\nu(F_x)}\mathbf{1}_{\{z\in F_x\}},$$

which implies that, almost surely,

$$\mathbf{P}_{\mathcal{H}}(Z_1\in\mathrm{Sub}(\mathcal{H},z)\,|\,Z_0)=\sum_{x\in[\![Y_0,Z_0]\!]}\frac{\nu(\mathrm{Sub}(\mathcal{H},z))}{\nu(F_x)}\mathbf{1}_{\{z\in F_x\}}\nu(F_x)=\nu(\mathrm{Sub}(\mathcal{H},z)),$$

so that Z_1 has distribution ν. For the induction step, suppose now that Z_k has distribution ν. Conditionally on Y_k and Z_{k-1}, the point Z_k is distributed according to the restriction of ν to the set $\mathrm{Fr}(\mathcal{H},Y_k,Z_{k-1})$. Applying the result for $k=1$ to the subtree $\mathrm{Fr}(\mathcal{H},Y_k,Z_{k-1})$, we see that Z_{k+1} is also distributed according to the restriction of ν to $\mathrm{Fr}(\mathcal{H},Y_k,Z_{k-1})$. So, Z_{k+1} and Z_k have the same conditional distribution, and it follows that Z_{k+1} is ν-distributed. Finally, since $Z_k\to U$ almost surely and for every k, Z_k is ν-distributed, it follows that U is a ν-distributed leaf and the proof is complete. □

Finally, we prove that U does not contain any of the "auxiliary" randomness used for the construction in the following sense:

Lemma 5.19. *Conditionally on $\mathcal{H}$, the random leaf U is a measurable function of $(A_x,x\in\mathrm{Br}(\mathcal{H}))$.*

Proof. It is clear from the construction, that U is a measurable function of $(A_x,x\in\mathrm{Br}(\mathcal{H}))$ and Z_0. It thus suffices to show that U is independent of Z_0. To see this, consider an independent copy Z_0' of Z_0 and let $(Y_k',Z_k')_{k\ge1}$ be the sequence of random variables obtained from this initial choice. By Lemma 5.18, Z_k' converges almost surely to a leaf that we denote by U'. We now show that a.s. $U'=U$.

In this direction, we prove by induction that for every $k\ge1$ we have $Y_k'\in[\![\rho,U]\!]$ and $Z_k'\in\mathrm{Fr}(\mathcal{H},Y_k',U)$. To see that this is the case, observe that with probability one $Z_0'\wedge U\in\mathrm{Sk}(\mathcal{H})$ is a turning point for Z_0'. It is also the closest such point from the root and thus $Y_1'=Z_0'\wedge U$. Moreover, since $(Y_k)_{k\ge0}$ is a path to U that is consistent with $(A_x,x\in\mathrm{Br}(\mathcal{H}))$, by construction $U\in\mathrm{Sub}(\mathcal{H},Y_1')\setminus\mathrm{Fr}(\mathcal{H},Y_1',Z_1')$. In other words, $Z_1'\in\mathrm{Fr}(\mathcal{H},Y_1',U)$. Supposing now that the claim holds for all integers up to $k\ge1$, we see that $Z_k'\wedge U\in]\!]Y_k',U]\!]$ since $\mathcal{H}$ is a.s. binary. Again, $Z_k'\wedge U$ is a turning point, and there is no other such point on $]\!]Y_k',Z_k'\wedge U]\!]$ so that $Y_{k+1}'=Z_k'\wedge U$. As before, $Y_{k+1}'\in[\![Y_r,Y_{r+1}]\!]$ for some $r\ge0$, and because the path to U is consistent, it must be the case that $Z_k'\in\mathrm{Fr}(\mathcal{H},Y_{k+1}',U)$.

Finally, recall that we proved in Lemma 5.18 that Y_k' also almost surely converges to U'. Since $U\in\mathrm{Sub}(\mathcal{H},Y_k')$ for each $k\ge1$, we have $U=U'$ and the proof is complete. □

5.4.2 The direct shuffle as the limit of k-reversals

It is now easy to use Lemma 5.18 in order to construct a sequence of i.i.d. leaves $(U_i)_{i\ge1}$ which are distributed according to the mass measure ν, and is consistent with the collection $(A_x,x\in\mathrm{Br}(\mathcal{H}))$. We proceed inductively as follows. First set U_1 to be the ν-distributed leaf whose existence is guaranteed by Lemma 5.18. Then, assume that we have defined $(U_i)_{1\le i\le k}$ and set $\mathcal{S}_k=\cup_{1\le i\le k}[\![\rho,U_i]\!]$. Let U_{k+1}° be an independent ν-leaf in $\mathcal{H}$. With probability one $U_{k+1}^\circ\notin\mathcal{S}_k$, so $R_{k+1}:=\{s\in\mathcal{H}:[\![s,U_{k+1}^\circ]\!]\cap\mathcal{S}_k=\varnothing\}$ is a non-empty subtree of $\mathcal{H}$ that also has positive mass. Then U_{k+1}° is distributed according to the restriction of ν to R_{k+1}, but may not be consistent with $(A_x,x\in\mathrm{Br}(R_{k+1}))$ in that subtree. By Lemma 5.18, there exists a random leaf U_{k+1}, distributed according to μ restricted to R_{k+1}, and which is consistent with the collection $(A_x, x \in \mathrm{Br}(R_{k+1}))$. One easily verifies that the collection $(U_i)_{i\geq 1}$ has the required properties.

Finally, the following result justifies the definition of $\mathrm{shuff}(\mathcal{H})$ that we gave in the introduction.

Proposition 5.20. *Let $\mathcal{H}$ be a Brownian CRT with mass measure ν. Let $(A_x, x \in \mathrm{Br}(\mathcal{H}))$ be a family of independent random variables such that A_x is distributed according to the restriction of ν to $\mathrm{Sub}(\mathcal{H}, x)$. Let $(U_i)_{i\geq 1}$ be a family of random leaves consistent with $(A_x, x \in \mathrm{Br}(\mathcal{H}))$. Then, the sequence*

$$(\mathrm{shuff}(\mathcal{H}; U_1, U_2, \dots, U_k))_{k\geq 1}$$

converges a.s. as $k \to \infty$ in the sense of Gromov–Prokhorov, and we define $\mathrm{shuff}(\mathcal{H})$ to be the limit.

Note that this definition is indeed consistent with the algorithm given in the introduction: for every fixed $k \geq 1$, the branch points on $\mathrm{Span}(\mathcal{H}; U_1, \dots, U_k)$ are used to form $\mathrm{shuff}(\mathcal{H}, U_1, \dots, U_k)$; furthermore, since every branch point $x \in \mathrm{Br}(\mathcal{H})$ is in some $\mathrm{Span}(\mathcal{H}; U_1, \dots, U_k)$ for k large enough, all the branch points are used to form the limit of $(\mathrm{shuff}(\mathcal{H}, U_1, \dots, U_k))_{k\geq 1}$.

5.5 Appendix: some facts about the Brownian CRT

In this section, we provide a proof of Lemma 5.12, which says that the sequence of (mass rescaled) real trees encountered on the path between two random points in the k-reveral are independent Brownian CRT. The proof is based on the scaling property of the Brownian excursion, Bismut's path decomposition of a Brownian excursion [33], and the size-biased sampling of a Poisson point process due to Perman et al. [94]. For the sake of precision, the arguments are phrased in terms of excursions rather than real trees, but we make sure to give the tree-based intuition as well.

BISMUT'S DECOMPOSITION. Let $B = (B_t)_{t\geq 0}$ be a standard Brownian motion. Let $\mathbf{N}$ be the Itô measure for the excursions of $|B|$ away from 0, which is a σ-finite measure on the space of non-negative continuous paths $\mathscr{C}(\mathbb{R}_+, \mathbb{R}_+)$. Let $w = (w_s)_{s\geq 0}$ be the coordinate process. In particular, if we denote by $\zeta := \inf\{s > 0 : w_s = 0\}$ the lifetime of an excursion, then $\mathbf{N}(\zeta \in dr) = dr/\sqrt{2\pi r^3}$. We denote by $\mathbf{N}^{(1)}$ the law of the normalized excursion. For $r > 0$, we set $\mathbf{N}^{(r)}$ to be the distribution of the rescaled process $(\sqrt{r} w_{s/r})_{0\leq s\leq r}$ under $\mathbf{N}^{(1)}$. Then it follows from the scaling property of Brownian motion that $\mathbf{N}^{(r)}$ is the law of w under $\mathbf{N}(\,\cdot\,|\,\zeta = r)$. We have

$$\mathbf{N}(\,\cdot\,) = \int \mathbf{N}(\zeta \in dr)\mathbf{N}^{(r)}(\,\cdot\,). \tag{5.28}$$

If we set $\mathscr{S}(w) \in \mathscr{C}(\mathbb{R}_+, \mathbb{R}_+)$ to be the process $(\mathscr{S}(w))(s) = \zeta^{-1/2} w_{s\zeta}$ for $s \geq 0$, then it follows from (5.28) that, under $\mathbf{N}$, $\mathscr{S}(w)$ is independent of ζ and is distributed as the normalized excursion.

Let $(Z_k)_{k\geq 0}$ be a sequence of independent variables uniformly distributed on $[0, \zeta]$. If we see w as encoding a Brownian CRT, then $(Z_k)_{k\geq 0}$ is a sequence of leaves sampled according to the mass measure. The CRT is decomposed along the path leading to the leaf corresponding to Z_0: set $K := w_{Z_0}$ and let

$$\overleftarrow{w}^{\circ}(t) = w_{(Z_0 - t)\vee 0}; \quad \overrightarrow{w}^{\circ}(t) = w_{t+Z_0}, \quad t \geq 0.$$

We need the following notation to describe precisely the spinal decomposition. Let $h : \mathbb{R}_+ \to \mathbb{R}_+$ be a continuous function with compact support and suppose that $h(0) > 0$. Define $\underline{h}(s) = \inf_{0\leq u\leq s} h(u)$. Let $\{(l_i, r_i), i \in \mathcal{I}(h)\}$ be the excursion intervals of $h - \underline{h}$ away from 0, which are the connected components of the open set $\{s \in \mathbb{R}_+ : h(s) - \underline{h}(s) > 0\}$. For each $i \in \mathcal{I}(h)$, let

$$h^i(s) = \big(h - \underline{h}\big)\big((l_i + s) \wedge r_i\big), \quad s \geq 0$$

be the excursion of $h - \underline{h}$ over the interval (l_i, r_i). We denote by $\mathcal{P}(h)$ the point measure on $\mathbb{R}_+ \times \mathscr{C}(\mathbb{R}_+, \mathbb{R}_+)$ defined by

$$\mathcal{P}(h) = \sum_{i \in \mathcal{I}(h)} \delta_{(h(0) - h(l_i),\, h^i)}.$$

We set $\mathcal{Q}$ to be the sum of the point measures $\mathcal{P}(\overrightarrow{w}^\circ)$ and $\mathcal{P}(\overleftarrow{w}^\circ)$ on $\mathbb{R}_+ \times \mathscr{C}(\mathbb{R}_+, \mathbb{R}_+)$:

$$\mathcal{Q} = \mathcal{P}(\overrightarrow{w}^\circ) + \mathcal{P}(\overleftarrow{w}^\circ) := \sum_{j \geq 1} \delta_{(s_j,\, w^j)},$$

where the last expression above serves as the definition of s_j and w^j, for $j \geq 1$.

Lemma 5.21 (Bismut's decomposition, [33])**.** *For each $m > 0$, $\mathcal{Q}$ under $\mathbf{N}(\,\cdot\,|K = m)$ has the distribution of a Poisson point measure of intensity measure $\mathbf{1}_{[0,m]} 2dt \otimes \mathbf{N}$.*

SIZE-BIASED ORDERING OF THE EXCURSIONS. For each $j \geq 1$, we denote by I_j the excursion interval (l_i, r_i) associated with w^j. We write ζ_j for the length of I_j. Notice that $\mathbf{N}$-a.s. $\sum_{j \geq 1} \zeta_j = \zeta$. Let $(\kappa_i)_{i \geq 1}$ be the permutation of $\mathbb{N}$ induced by $(Z_n)_{n \geq 1}$ as follows. Let κ_1 be the index such that $Z_1 \in I_{\kappa_1}$, then κ_1 is well-defined almost surely. For each $i \geq 1$, let $\sigma(i) := \inf\{n \geq 1 : Z_n \notin \cup_{1 \leq n \leq i} I_{\kappa_i}\}$, then let κ_{i+1} be the index such that $Z_{\sigma(i)} \in I_{\kappa_{i+1}}$. Let us point out that according to Lemma 5.21, $(\zeta_j)_{j \geq 1}$ under $\mathbf{N}(\,\cdot\,|\, K = m)$ is distributed as the jumps of a $1/2$-stable subordinator before time $2m$, and $(\zeta_{\kappa_i})_{i \geq 1}$ is the size-biased sampling introduced in [94]. Then by Palm's formula and the product form of the intensity measure in Lemma 5.21, we have

Lemma 5.22. *Under $\mathbf{N}^{(1)}(\,\cdot\,|\, K = m)$, the three sequences $(\zeta_{\kappa_i})_{i \geq 1}$, $(s_{\kappa_i})_{i \geq 1}$, $(\mathscr{S}(w^{\kappa_i}))_{i \geq 1}$ are independent. Moreover, $(\zeta_{\kappa_i})_{i \geq 1}$ is a Markov chain, $(s_{\kappa_i})_{i \geq 1}$ is an i.i.d. sequence of uniform variables on $[0, m]$, and $(\mathscr{S}(w^{\kappa_i}))_{i \geq 1}$ is an i.i.d. sequence of common law $\mathbf{N}^{(1)}$.*

Proof. Let $f, g, h : \mathbb{R}_+ \to \mathbb{R}$ be some continuous bounded functions, let $H : \mathscr{C}(\mathbb{R}_+, \mathbb{R}_+) \to \mathbb{R}$ and $G : \mathbb{R}_+ \times \mathscr{C}(\mathbb{R}_+, \mathbb{R}_+) \to \mathbb{R}$ be continuous and bounded. We use the notations:

$$\mathcal{Q}^- = \mathcal{Q} - \delta_{(s_{\kappa_1}, w^{\kappa_1})} \qquad \text{and} \qquad \mathcal{Q}^{j-} = \mathcal{Q} - \delta_{(s_j, w^j)}, \quad j \geq 1.$$

Note that $\zeta = \sum_{j \geq 1} \zeta_j$ is distributed as X_{2m}, where $(X_s)_{s \geq 0}$ is a $1/2$-stable subordinator. We denote by θ the density of X_{2m}. Then, by Palm's formula,

$$\begin{aligned}
&\mathbb{E}\left[f(s_{\kappa_1}) g(\zeta_{\kappa_1}) H(w^{\kappa_1}) G(\mathcal{Q}^-) h(\zeta)\right] \\
&= \mathbb{E}\left[\sum_{j \geq 1} \frac{\zeta_j}{\zeta} f(s_j) g(\zeta_j) H(w^j) G(\mathcal{Q}^{j-}) h(\zeta)\right] \\
&= \int_0^\infty dz \theta(z) h(r+z) \int_0^m 2dt g(t) \int_0^\infty \frac{dr}{\sqrt{2\pi r^3}} f(r) \frac{r}{r+z} \mathbf{N}^{(r)}\big(H(w^j)\big) \mathbb{E}[G(\mathcal{Q}) \,|\, \zeta = z].
\end{aligned}$$

It follows that, for $0 < r < 1$, and $0 < t < m$, we have

$$\mathbf{P}\left(\zeta_{\kappa_1} \in dr, s_{\kappa_1} \in dt, \mathscr{S}(w^{\kappa_1}) \in dw \,\middle|\, \zeta = 1\right) = \frac{\theta(1-r)dr}{\Lambda\sqrt{2\pi r}} \cdot \frac{dt}{m} \cdot \mathbf{N}^{(1)}(dw)$$

where $\Lambda = \int_0^1 dr \theta(1-r)/\sqrt{2\pi r}$, and given $\zeta_{\kappa_1} = r$, $\mathcal{Q}^-$ is independent of $(s_{\kappa_1}, \mathscr{S}(w^{\kappa_1}))$ and has distribution $\mathbf{N}(\,\cdot\,|\, \zeta = 1-r, K = m)$. From there, a simple induction argument yields the claim. $\square$

THE TREES ALONG THE PATH BETWEEN TWO RANDOM POINTS. We are finally in position of proving Lemma 5.12. Let $(\pi_j)_{j\geq 1}$ be the subsequence of $(\kappa_i)_{i\geq 1}$ defined by

$$\pi_1 = \kappa_1, \quad \text{and if } \tau_j = \inf\{i > \pi_j : s_{\kappa_i} > s_{\pi_j}\} \text{ then } \pi_{j+1} = \kappa_{\tau_j}, \quad j \geq 1.$$

Then by definition, $(s_{\pi_j}, j \geq 1)$ has the same distribution as $(d(\rho, x_1(m)), m \geq 0)$, where $(x_1(m), m \geq 0)$ is the sequence defined in Proposition 5.6. Observe that $(\pi_j)_{j\geq 1}$ depends only on $(s_{\kappa_i})_{i\geq 1}$, which is independent of $(\mathscr{S}(w^{\kappa_i}))_{i\geq 1}$ according to the previous lemma. We deduce that

Lemma 5.23. *Under $\mathbf{N}^{(1)}(\,\cdot\,|\,K = m)$, the three sequences $(\zeta_{\pi_j})_{j\geq 1}$, $(s_{\pi_j})_{j\geq 1}$, and $(\mathscr{S}(w^{\pi_j}))_{j\geq 1}$ are independent. Moreover, $(\mathscr{S}(w^{\pi_j}), j \geq 1)$ is an i.i.d. sequence of common law $\mathbf{N}^{(1)}$.*

Finally, let $(Z'_k)_{k\geq 1}$ be another sequence of i.i.d. random variables uniformly distributed on $[0, \zeta]$, independent of $(Z_k)_{k\geq 1}$. Define $(\kappa'_i)_{i\geq 1}$ and $(\pi'_j)_{j\geq 1}$ using $(Z'_k)_{k\geq 1}$ in the same way that $(\kappa_i)_{i\geq 1}$ and $(\pi_j)_{j\geq 1}$ were defined using $(Z_k)_{k\geq 1}$. Let

$$\mathscr{I} := \inf\{j \geq 1 : \exists k \geq 1 \text{ such that } s_{\kappa_j} = s_{\kappa'_k}\}, \quad \mathscr{I}' := \inf\{k \geq 1 : \exists j \geq 1 \text{ such that } s_{\kappa'_k} = s_{\kappa_j}\}.$$

Note that for each pair (j_0, k_0), the event $\{\mathscr{I} = j_0, \mathscr{I}' = k_0\}$ depends only on $(\zeta_{\pi_j}, 1 \leq j \leq j_0)$ and $(\zeta_{\pi'_k}, 1 \leq k \leq k_0)$. Therefore, on this event, $\{\mathscr{S}(w^{\pi_j}), 1 \leq j \leq j_0\} \cup \{\mathscr{S}(w^{\pi_k}), 1 \leq k \leq k_0 - 1\}$ are $j_0 + k_0 - 1$ independent copies of w under $\mathbf{N}^{(1)}$. Integrating with respect to the law of K under $\mathbf{N}^{(1)}$ (the Rayleigh distribution) proves Lemma 5.12.

Bibliography

[1] R. Abraham and J.-F. Delmas. Asymptotics for the small fragments of the fragmentation at nodes. *Bernoulli*, 13(1):211–228, 2007.

[2] R. Abraham and J.-F. Delmas. Fragmentation associated with Lévy processes using snake. *Probab. Theory Related Fields*, 141(1-2):113–154, 2008.

[3] R. Abraham and J.-F. Delmas. Williams' decomposition of the Lévy continuum random tree and simultaneous extinction probability for populations with neutral mutations. *Stochastic Process. Appl.*, 119(4): 1124–1143, 2009.

[4] R. Abraham and J.-F. Delmas. Record process on the continuum random tree. *ALEA Lat. Am. J. Probab. Math. Stat.*, 10(1):225–251, 2013.

[5] R. Abraham and J.-F. Delmas. The forest associated with the record process on a Lévy tree. *Stochastic Process. Appl.*, 123(9):3497–3517, 2013.

[6] R. Abraham, J.-F. Delmas, and G. Voisin. Pruning a Lévy continuum random tree. *Electron. J. Probab.*, 15: no. 46, 1429–1473, 2010.

[7] L. Addario-Berry, N. Broutin, and C. Holmgren. Cutting down trees with a Markov chainsaw. *The Annals of Applied Probability*, 2013. (To appear).

[8] D. Aldous. The continuum random tree. I. *Ann. Probab.*, 19(1):1–28, 1991.

[9] D. Aldous. The continuum random tree. II. An overview. In *Stochastic analysis (Durham, 1990)*, volume 167 of *London Math. Soc. Lecture Note Ser.*, pages 23–70. Cambridge Univ. Press, Cambridge, 1991.

[10] D. Aldous. The continuum random tree. III. *Ann. Probab.*, 21(1):248–289, 1993.

[11] D. Aldous and J. Pitman. The standard additive coalescent. *Ann. Probab.*, 26(4):1703–1726, 1998.

[12] D. Aldous and J. Pitman. A family of random trees with random edge lengths. *Random Structures Algorithms*, 15(2):176–195, 1999.

[13] D. Aldous and J. Pitman. Inhomogeneous continuum random trees and the entrance boundary of the additive coalescent. *Probab. Theory Related Fields*, 118(4):455–482, 2000.

[14] D. Aldous and J. Pitman. Invariance principles for non-uniform random mappings and trees. In *Asymptotic combinatorics with application to mathematical physics (St. Petersburg, 2001)*, volume 77 of *NATO Sci. Ser. II Math. Phys. Chem.*, pages 113–147. Kluwer Acad. Publ., Dordrecht, 2002.

[15] D. Aldous, G. Miermont, and J. Pitman. The exploration process of inhomogeneous continuum random trees, and an extension of Jeulin's local time identity. *Probab. Theory Related Fields*, 129(2):182–218, 2004.

[16] D. Aldous, G. Miermont, and J. Pitman. Brownian bridge asymptotics for random p-mappings. *Electron. J. Probab.*, 9:no. 3, 37–56 (electronic), 2004.

[17] D. Aldous, G. Miermont, and J. Pitman. Weak convergence of random p-mappings and the exploration process of inhomogeneous continuum random trees. *Probab. Theory Related Fields*, 133(1):1–17, 2005.

[18] D. J. Aldous. The random walk construction of uniform spanning trees and uniform labelled trees. *SIAM J. Discrete Math.*, 3(4):450–465, 1990.

[19] V. Anantharam and P. Tsoucas. A proof of the Markov chain tree theorem. *Statist. Probab. Lett.*, 8(2): 189–192, 1989.

[20] K. B. Athreya and P. E. Ney. *Branching processes*. Springer-Verlag, New York-Heidelberg, 1972. Die

Grundlehren der mathematischen Wissenschaften, Band 196.

[21] E. Baur and J. Bertoin. Cutting edges at random in large recursive trees. hal-00982497, 2014.

[22] J. Bertoin. *Lévy processes*, volume 121 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1996.

[23] J. Bertoin. A fragmentation process connected to Brownian motion. *Probab. Theory Related Fields*, 117(2): 289–301, 2000.

[24] J. Bertoin. Eternal additive coalescents and certain bridges with exchangeable increments. *Ann. Probab.*, 29 (1):344–360, 2001.

[25] J. Bertoin. Self-similar fragmentations. *Ann. Inst. H. Poincaré Probab. Statist.*, 38(3):319–340, 2002.

[26] J. Bertoin. *Random fragmentation and coagulation processes*, volume 102 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2006.

[27] J. Bertoin. Fires on trees. *Ann. Inst. Henri Poincaré Probab. Stat.*, 48(4):909–921, 2012.

[28] J. Bertoin. The cut-tree of large recursive trees. *Ann. Inst. Henri Poincaré Probab. Stat.*, 2013. (to appear).

[29] J. Bertoin and A. V. Gnedin. Asymptotic laws for nonconservative self-similar fragmentations. *Electron. J. Probab.*, 9:no. 19, 575–593, 2004.

[30] J. Bertoin and G. Miermont. The cut-tree of large Galton-Watson trees and the Brownian CRT. *Ann. Appl. Probab.*, 23(4):1469–1493, 2013.

[31] P. Billingsley. *Convergence of probability measures*. John Wiley & Sons Inc., New York, 1968.

[32] N. H. Bingham. Continuous branching processes and spectral positivity. *Stochastic Processes Appl.*, 4(3): 217–242, 1976.

[33] J.-M. Bismut. Last exit decompositions and regularity at the boundary of transition probabilities. *Z. Wahrsch. Verw. Gebiete*, 69(1):65–98, 1985.

[34] R. M. Blumenthal. *Excursions of Markov processes*. Probability and its Applications. Birkhäuser Boston Inc., Boston, MA, 1992.

[35] A. N. Borodin and P. Salminen. *Handbook of Brownian motion—facts and formulae*. Probability and its Applications. Birkhäuser Verlag, Basel, 1996.

[36] S. Boucheron, G. Lugosi, and P. Massart. *Concentration inequalities*. Oxford University Press, Oxford, 2013. A nonasymptotic theory of independence, With a foreword by Michel Ledoux.

[37] A. Broder. Generating random spanning trees. In *Proc. 30'th IEEE Symp. Found. Comp. Sci.*, pages 442–447. 1989.

[38] N. Broutin and P. Flajolet. The distribution of height and diameter in random non-plane binary trees. *Random Structures Algorithms*, 41(2):215–252, 2012.

[39] N. Broutin and M. Wang. Cutting down p-trees and inhomogeneous continuum random trees. arXiv:1408.0144 [math.PR], 2014.

[40] N. Broutin and M. Wang. Reversing the cut tree of the Brownian CRT. arXiv:1408.2924 [math.PR], 2014.

[41] M. Camarri and J. Pitman. Limit distributions and random trees derived from the birthday problem with unequal probabilities. *Electron. J. Probab.*, 5:no. 2, 18 pp. (electronic), 2000.

[42] A. Cayley. A theorem on trees. *Quarterly Journal of Pure and Applied Mathematics*, 23:376–378, 1889.

[43] J. M. Chambers, C. L. Mallows, and B. W. Stuck. A method for simulating stable random variables. *J. Amer. Statist. Assoc.*, 71(354):340–344, 1976.

[44] L. Chaumont. Excursion normalisée, méandre et pont pour les processus de Lévy stables. *Bull. Sci. Math.*, 121(5):377–403, 1997.

[45] K. L. Chung. Excursions in Brownian motion. *Ark. Mat.*, 14(2):155–177, 1976.

[46] D. Dieuleveut. The vertex-cut-tree of Galton-Watson trees converging to a stable tree. arXiv:1312.5525 [math.PR], 2013.

[47] A. Dress, V. Moulton, and W. Terhalle. T-theory: an overview. *European J. Combin.*, 17(2-3):161–175, 1996. Discrete metric spaces (Bielefeld, 1994).

[48] M. Drmota, A. Iksanov, M. Moehle, and U. Roesler. A limiting distribution for the number of cuts needed

to isolate the root of a random recursive tree. *Random Structures Algorithms*, 34(3):319–336, 2009.

[49] T. Duquesne. A limit theorem for the contour process of conditioned Galton-Watson trees. *Ann. Probab.*, 31(2):996–1027, 2003.

[50] T. Duquesne. The coding compact real trees by real valued functions. arXiv:math/0604106 [math.PR], 2006.

[51] T. Duquesne and J.-F. Le Gall. Random trees, Lévy processes and spatial branching processes. *Astérisque*, (281):vi+147, 2002.

[52] T. Duquesne and J.-F. Le Gall. Probabilistic and fractal aspects of Lévy trees. *Probab. Theory Related Fields*, 131(4):553–603, 2005.

[53] T. Duquesne and J.-F. Le Gall. On the re-rooting invariance property of Lévy trees. *Electron. Commun. Probab.*, 14:317–326, 2009.

[54] T. Duquesne and M. Wang. Decomposition of Lévy trees along their diameter. 2014.

[55] T. Duquesne and M. Winkel. Growth of Lévy trees. *Probab. Theory Related Fields*, 139(3-4):313–371, 2007.

[56] M. Dwass. Branching processes in simple random walk. *Proc. Amer. Math. Soc.*, 51:270–274, 1975.

[57] S. N. Evans. *Probability and real trees*, volume 1920 of *Lecture Notes in Mathematics*. Springer, Berlin, 2008. Lectures from the 35th Summer School on Probability Theory held in Saint-Flour, July 6–23, 2005.

[58] S. N. Evans and J. Pitman. Construction of Markovian coalescents. *Ann. Inst. H. Poincaré Probab. Statist.*, 34(3):339–383, 1998.

[59] S. N. Evans, J. Pitman, and A. Winter. Rayleigh processes, real trees, and root growth with re-grafting. *Probab. Theory Related Fields*, 134(1):81–126, 2006.

[60] J. A. Fill, N. Kapur, and A. Panholzer. Destruction of very simple trees. *Algorithmica*, 46(3-4):345–366, 2006.

[61] C. Goldschmidt and B. Haas. Behavior near the extinction time in self-similar fragmentations. I. The stable case. *Ann. Inst. Henri Poincaré Probab. Stat.*, 46(2):338–368, 2010.

[62] A. Greven, P. Pfaffelhuber, and A. Winter. Convergence in distribution of random metric measure spaces (Λ-coalescent measure trees). *Probab. Theory Related Fields*, 145(1-2):285–322, 2009.

[63] A. Grimvall. On the convergence of sequences of branching processes. *Ann. Probability*, 2:1027–1045, 1974.

[64] M. Gromov. *Metric structures for Riemannian and non-Riemannian spaces*. Modern Birkhäuser Classics. Birkhäuser Boston Inc., Boston, MA, english edition, 2007.

[65] B. Haas and G. Miermont. The genealogy of self-similar fragmentations with negative index as a continuum random tree. *Electron. J. Probab.*, 9:no. 4, 57–97 (electronic), 2004.

[66] B. Haas and G. Miermont. Scaling limits of Markov branching trees with applications to Galton-Watson and random unordered trees. *Ann. Probab.*, 40(6):2589–2666, 2012.

[67] C. Holmgren. Random records and cuttings in binary search trees. *Combin. Probab. Comput.*, 19(3):391–424, 2010.

[68] C. Holmgren. A weakly 1-stable distribution for the number of random records and cuttings in split trees. *Adv. in Appl. Probab.*, 43(1):151–177, 2011.

[69] I. A. Ibragimov and K. E. Černin. On the unimodality of stable laws. *Teor. Veroyatnost. i Primenen.*, 4: 453–456, 1959.

[70] A. Iksanov and M. Möhle. A probabilistic proof of a weak limit law for the number of cuts needed to isolate the root of a random recursive tree. *Electron. Comm. Probab.*, 12:28–35, 2007.

[71] J. Jacod and A. N. Shiryaev. *Limit theorems for stochastic processes*, volume 288 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, Berlin, 1987.

[72] S. Janson. Random cutting and records in deterministic and random trees. *Random Structures Algorithms*, 29(2):139–179, 2006.

[73] M. Jiřina. Stochastic branching processes with continuous state space. *Czechoslovak Math. J.*, 8 (83):

292–313, 1958.

[74] O. Kallenberg. Canonical representations and convergence criteria for processes with interchangeable increments. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 27:23–36, 1973.

[75] J. F. C. Kingman. The coalescent. *Stochastic Process. Appl.*, 13(3):235–248, 1982.

[76] I. Kortchemski. Invariance principles for Galton-Watson trees conditioned on the number of leaves. *Stochastic Process. Appl.*, 122(9):3126–3172, 2012.

[77] A. E. Kyprianou. *Introductory lectures on fluctuations of Lévy processes with applications*. Universitext. Springer-Verlag, Berlin, 2006.

[78] J. Lamperti. The limit of a sequence of branching processes. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 7:271–288, 1967.

[79] J. Lamperti. Continuous state branching processes. *Bull. Amer. Math. Soc.*, 73:382–386, 1967.

[80] J.-F. Le Gall. The uniform random tree in a Brownian excursion. *Probab. Theory Related Fields*, 96(3):369–383, 1993.

[81] J.-F. Le Gall. *Spatial branching processes, random snakes and partial differential equations*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 1999.

[82] J.-F. Le Gall. Random trees and applications. *Probab. Surv.*, 2:245–311, 2005.

[83] J.-F. Le Gall and Y. Le Jan. Branching processes in Lévy processes: the exploration process. *Ann. Probab.*, 26(1):213–252, 1998.

[84] W. Löhr. Equivalence of Gromov-Prohorov- and Gromov's $\underline{\Box}_\lambda$-metric on the space of metric measure spaces. *Electron. Commun. Probab.*, 18:no. 17, 10, 2013.

[85] W. Löhr, G. Voisin, and A. Winter. Convergence of bi-measure $\mathbb{R}$-Trees and the pruning process. arxiv:1304.6035 [math.PR], 2013.

[86] R. Lyons and Y. Peres. *Probability on Trees and Networks*. Cambridge University Press, 2014. In preparation. Current version available at http://mypage.iu.edu/~rdlyons/.

[87] P. Marchal. A note on the fragmentation of a stable tree. In *Fifth Colloquium on Mathematics and Computer Science*, Discrete Math. Theor. Comput. Sci. Proc., AI, pages 489–499. Assoc. Discrete Math. Theor. Comput. Sci., Nancy, 2008.

[88] A. Meir and J. W. Moon. Cutting down random trees. *J. Austral. Math. Soc.*, 11:313–324, 1970.

[89] G. Miermont. Self-similar fragmentations derived from the stable tree. I. Splitting at heights. *Probab. Theory Related Fields*, 127(3):423–454, 2003.

[90] G. Miermont. Self-similar fragmentations derived from the stable tree. II. Splitting at nodes. *Probab. Theory Related Fields*, 131(3):341–375, 2005.

[91] G. Miermont. Tessellations of random maps of arbitrary genus. *Ann. Sci. Éc. Norm. Supér. (4)*, 42(5):725–781, 2009.

[92] J. Neveu. Arbres et processus de Galton-Watson. *Ann. Inst. H. Poincaré Probab. Statist.*, 22(2):199–207, 1986.

[93] A. Panholzer. Cutting down very simple trees. *Quaest. Math.*, 29(2):211–227, 2006.

[94] M. Perman, J. Pitman, and M. Yor. Size-biased sampling of Poisson point processes and excursions. *Probab. Theory Related Fields*, 92(1):21–39, 1992.

[95] J. Pitman. *Combinatorial stochastic processes*, volume 1875 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2006. Lectures from the 32nd Summer School on Probability Theory held in Saint-Flour, July 7–24, 2002, With a foreword by Jean Picard.

[96] A. Rényi. On the enumeration of trees. In *Combinatorial Structures and their Applications (Proc. Calgary Internat. Conf., Calgary, Alta., 1969)*, pages 355–360. Gordon and Breach, New York, 1970.

[97] D. Revuz and M. Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, third edition, 1999.

[98] G. Szekeres. Distribution of labelled trees by diameter. In *Combinatorial mathematics, X (Adelaide, 1982)*,

volume 1036 of *Lecture Notes in Math.*, pages 392–397. Springer, Berlin, 1983.

[99] W. Vervaat. A relation between Brownian bridge and Brownian excursion. *Ann. Probab.*, 7(1):143–149, 1979.

[100] M. Wang. Height and diameter of Brownian tree. 2014.

[101] A. Weil. *Elliptic functions according to Eisenstein and Kronecker*. Springer-Verlag, Berlin-New York, 1976. Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 88.

[102] E. T. Whittaker and G. N. Watson. *A course of modern analysis. An introduction to the general theory of infinite processes and of analytic functions: with an account of the principal transcendental functions*. Fourth edition. Reprinted. Cambridge University Press, New York, 1962.

[103] V. M. Zolotarev. *One-dimensional stable distributions*, volume 65 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 1986. Translated from the Russian by H. H. McFaden, Translation edited by Ben Silver.